

(19) 日本国特許庁 (J P)

(12) 特 許 公 報 (B 2)

(11) 特許番号

特許第3280211号
(P3280211)

(45) 発行日 平成14年4月30日 (2002. 4. 30)

(24) 登録日 平成14年2月22日 (2002. 2. 22)

(51) Int.Cl.⁷

H 0 4 N 7/32

識別記号

F I

H 0 4 N 7/137

Z

請求項の数30(全 16 頁)

(21) 出願番号 特願平7-293835

(22) 出願日 平成7年10月18日 (1995. 10. 18)

(65) 公開番号 特開平8-214305

(43) 公開日 平成8年8月20日 (1996. 8. 20)

審査請求日 平成10年3月10日 (1998. 3. 10)

(31) 優先権主張番号 3 2 7 5 3 0

(32) 優先日 平成6年10月21日 (1994. 10. 21)

(33) 優先権主張国 米国 (U S)

前置審査

(73) 特許権者 390035493

エイ・ティ・アンド・ティ・コーポレー
ション

A T & T C O R P .

アメリカ合衆国 10013-2412 ニュー
ヨーク ニューヨーク アヴェニュー
オブ ジ アメリカズ 32

(72) 発明者
ネルソン ボッツフォード ザ サード
アメリカ合衆国, 08861 ニュージャ
ージー, サマーヴィル, マグノリア レイ
ン 605

(74) 代理人 100081053

弁理士 三俣 弘文

審査官 古川 哲也

最終頁に続く

(54) 【発明の名称】 ビデオエンコーダ制御システム

1

(57) 【特許請求の範囲】

【請求項1】 ビデオエンコーダを有するビデオエン
コーダ制御システムにおいて、該ビデオエンコーダはプロ
セッサを有し、

前記ビデオエンコーダに入力される入力ビデオデー
タは、第1フレームを含む複数のフレームを有し、各フレ
ームは複数のフィールドを有し、各フィールドはパリティ
を有し、

前記プロセッサは、

N > 1 として、所定数N個のフレームだけ入力ビデオ
データを遅延させて遅延ビデオデータを生成する多フィー
ルド遅延回路と、

前記入力ビデオデータを処理して前記第1フレームの統
計値を生成する統計値生成器とを有し、

前記統計値は、ローパスフィルタイメージ値のサブサン

2

プリング値と、ローパスフィルタイメージ値のサブサン
プリング値のピクセル値のフィールド差の絶対値の1フ
ィールドにわたる最大値とを含み、

前記統計値生成器は、1フィールドのピクセルのブロッ
クの平均を、該フィールドのローパスフィルタイメージ
値のサブサンプリング値として生成し、

前記プロセッサはさらに、

フィルム検出を実行するために前記最大値を用いてノン
フィルム状態を検出する第1ディテクタを有することを
特徴とするビデオエンコーダ制御システム。

【請求項2】 前記プロセッサは、前記入力ビデオデー
タをM P E G 標準で符号化するように前記ビデオエン
コーダとともに動作することを特徴とする請求項1に記載
のビデオエンコーダ制御システム。

【請求項3】 前記統計値生成器は、同一パリティを有

する連続するフィールドのピクセル間のフィールド差の絶対値の和を計算し、
前記プロセッサは、
前記第 1 フレームから次のフレームへのシーン変化を検出してシーン変化フラグ信号を制御信号として生成する第 2 ディテクタと、
前記シーン変化フラグ信号を用いてエンコーダモジュールを制御するレートコントローラとをさらに有することを特徴とする請求項 1 に記載のビデオエンコーダ制御システム。

【請求項 4】 前記統計値生成器は、各フィールドの平均ピクセル値を計算し、
前記第 2 ディテクタは、前記第 1 フレームの第 1 フィールドと連続するフィールドとに関連する平均ピクセル値を用いてビデオフェードを決定することを特徴とする請求項 3 に記載のビデオエンコーダ制御システム。

【請求項 5】 前記統計値生成器は、同一パリティを有する連続するフィールドのピクセル間のフィールド差の絶対値の和を計算し、
前記ビデオエンコーダ制御システムは、
前記フィールド差の絶対値の和を用いて前記第 1 フレームの符号化の解像度を変化させる解像度セクタをさらに有することを特徴とする請求項 1 に記載のビデオエンコーダ制御システム。

【請求項 6】 前記解像度セクタは、前記和にตอบสนองして解像度選択信号を生成し、
前記プロセッサは、
前記解像度選択信号にตอบสนองして、該解像度選択信号に対応する解像度で前記遅延ビデオデータをサンプリングするサンプラをさらに有することを特徴とする請求項 5 に記載のビデオエンコーダ制御システム。

【請求項 7】 前記統計値生成器は、同一パリティを有する連続するフィールドのピクセル間のフィールド差の絶対値の和を計算し、
前記第 1 ディテクタは、前記和を用いて関連するフィールド内の冗長フィールドを検出し、該冗長フィールドに対応する冗長性フラグ信号を生成し、
前記プロセッサは、
前記冗長性フラグ信号を用いてエンコーダモジュールを制御するレートコントローラをさらに有することを特徴とする請求項 1 に記載のビデオエンコーダ制御システム。

【請求項 8】 前記第 1 ディテクタは、入力ビデオデータに対する 3 : 2 プルダウン検出を実行することを特徴とする請求項 7 に記載のビデオエンコーダ制御システム。

【請求項 9】 前記第 1 ディテクタは、前記 3 : 2 プルダウン検出にตอบสนองして、フィルムデータが前記入力ビデオデータにおいて検出されたかどうかを示すフィルムフラグ信号を生成し、

前記プロセッサは、
前記フィルムフラグ信号にตอบสนองして、複数のフィルタ値を用いて前記遅延ビデオデータをフィルタ処理するプレフィルタをさらに有することを特徴とする請求項 8 に記載のビデオエンコーダ制御システム。

【請求項 10】 前記プレフィルタは、フィールド差の絶対値の和とエンコーダビットレートとを用いて、前記遅延ビデオデータをフィルタ処理するために前記複数のフィルタ値のうちの少なくとも 1 つを選択することを特徴とする請求項 9 に記載のビデオエンコーダ制御システム。

【請求項 11】 前記第 1 ディテクタは、前記第 1 フレームと連続するフレームとに関連するフィールド差の絶対値の最大値を用いて、局所的に変化する領域を判定することを特徴とする請求項 7 に記載のビデオエンコーダ制御システム。

【請求項 12】 前記プロセッサは、
前記第 1 フレームから次のフレームへのシーン変化を検出してシーン変化フラグを生成する第 2 ディテクタと、
周期的カウントを生成するカウンタとを有し、
前記プロセッサは、前記周期的カウントにตอบสนองして内部フレームフラグを生成し、
前記プロセッサは、前記シーン変化フラグにตอบสนองして内部フレームフラグを生成し、
前記レートコントローラは、前記内部フレームフラグにตอบสนองして、前記複数のフレーム内のシーン変化に対応する位置に内部フレームを挿入することを特徴とする請求項 11 に記載のビデオエンコーダ制御システム。

【請求項 13】 前記プレフィルタは、
フィルタのオーバーフローを回避するために、フィルタ処理された遅延ビデオデータを所定範囲内にクランプするビデオデータクランプを有することを特徴とする請求項 12 に記載のビデオエンコーダ制御システム。

【請求項 14】 前記プロセッサは、前記シーン変化フラグにตอบสนองして前記カウンタのカウントを更新し、該カウンタの更新にตอบสนองして内部フレームフラグを生成することを特徴とする請求項 13 に記載のビデオエンコーダ制御システム。

【請求項 15】 符号化されるべき複数のフレームに対応する入力ビデオデータ内のフィルムおよびシーン変化を検出するビデオエンコーダにおいて、
前記ビデオエンコーダに入力される入力ビデオデータは、第 1 フレームを含む複数のフレームを有し、各フレームは複数のフィールドを有し、各フィールドはパリティを有し、
前記ビデオエンコーダは、
N > 1 として、所定数 N 個のフレームだけ入力ビデオデータを遅延させて第 1 遅延ビデオデータを生成する多フィールド遅延回路と、

ローパスフィルタを有し、前記入力ビデオデータを処理

して前記第1フレームの統計値を制御信号として生成する統計値生成器とを有し、

前記統計値は、ローパスフィルタイメージ値のサブサンプリング値と、ローパスフィルタイメージ値のサブサンプリング値のフィールド差の絶対値の和と、ローパスフィルタイメージ値のサブサンプリング値のピクセル値のフィールド差の絶対値の1フィールドにわたる最大値とを含み、

前記統計値生成器は、同一パリティを有する連続するフィールドのピクセル間のフィールド差の絶対値の和を計算し、 10

前記ビデオエンコーダはさらに、

前記統計値を用いて、前記最大値から前記第1フレームのノンフィルム状態を検出してフィルムデータが検出されたかどうかを示すフィルムフラグを生成し、前記絶対値の和が所定しきい値を超えたことから前記複数のフレーム間のシーン変化を検出してシーン変化フラグを生成し、前記複数のフレーム間のビデオフェード状態を検出してビデオフェードフラグを生成する第1ディテクタと、

エンコーダモジュールと、

生成されたフィルムフラグ、生成されたシーン変化フラグ、および生成されたビデオフェードフラグにตอบสนองして、少なくとも1つの制御信号を生成し、前記エンコーダモジュールが前記第1遅延ビデオデータを符号化するのを制御するために、該少なくとも1つの制御信号を前記エンコーダモジュールへ送信するレートコントローラとを有することを特徴とするビデオエンコーダ。

【請求項16】 前記ビデオエンコーダは、前記和にตอบสนองして解像度選択信号を生成する解像度セレクトをさらに有し、 30

前記レートコントローラは、前記解像度選択信号にตอบสนองして、前記第1フレームの符号化の解像度を変化させることを特徴とする請求項15に記載のビデオエンコーダ。

【請求項17】 前記ビデオエンコーダは、入力コマンドを処理する制御モジュールをさらに有し、

前記第1ディテクタは、予測されるフィルムコマンドにตอบสนองして、フィルムフラグを生成することを特徴とする請求項15に記載のビデオエンコーダ。 40

【請求項18】 前記統計値生成器は、同一パリティを有する連続するフィールドのピクセル間のフィールド差の絶対値の和を計算し、

前記第1ディテクタは、前記和を用いて前記第1フレームから次のフレームへのシーン変化を検出することを特徴とする請求項15に記載のビデオエンコーダ。

【請求項19】 前記統計値生成器は、各フィールドの平均ピクセル値を計算し、

前記ビデオエンコーダは、

前記第1フレームの第1フィールドと連続するフィール 50

ドに関連する平均ピクセル値を用いてビデオフェード状態を判定し、対応するビデオフェードフラグを生成する第2ディテクタをさらに有することを特徴とする請求項18に記載のビデオエンコーダ。

【請求項20】 前記統計値生成器は、同一パリティを有する連続するフィールドのピクセル間のフィールド差の絶対値の和を計算し、

前記第1ディテクタは、前記和を用いて関連するフィールドのうちの冗長フィールドを検出し、該冗長フィールドに対応してフィルムフラグを生成することを特徴とする請求項15に記載のビデオエンコーダ。

【請求項21】 前記多フィールド遅延回路は、前記入力ビデオデータを少なくとも1フィールドだけ遅延させて第2遅延ビデオデータを生成し、

前記統計値生成器は、前記入力ビデオデータと前記第2遅延ビデオデータからフィールド差の絶対値の和を生成することを特徴とする請求項20に記載のビデオエンコーダ。

【請求項22】 前記第1ディテクタは、前記第1フレームと連続するフレームとに関連するフィールド差の絶対値の最大値を用いて、局所的に変化する領域を判定することを特徴とする請求項20に記載のビデオエンコーダ。

【請求項23】 入力ビデオデータを符号化するビデオエンコーダを制御する方法において、

前記入力ビデオデータは、第1フレームを含む複数のフレームを有し、各フレームは複数のフィールドを有し、各フィールドはパリティを有し、

前記方法は、

$N > 1$ として、所定数 N 個のフレームだけ入力ビデオデータを遅延させて遅延ビデオデータを生成するステップと、

前記入力ビデオデータを処理して前記第1フレームを示すデータを生成する処理ステップとを有し、

前記データは、ローパスフィルタイメージ値のサブサンプリング値と、ローパスフィルタイメージ値のサブサンプリング値のフィールド差の絶対値の和と、ローパスフィルタイメージ値のサブサンプリング値のピクセル値のフィールド差の絶対値の1フィールドにわたる最大値とを含み、

前記方法はさらに、

前記入力ビデオデータ内にフィルムデータが検出されない場合に、フィルム検出のために前記最大値を用いて生成されるフィルムフラグ信号を生成するステップと、生成されたデータおよび前記フィルムフラグ信号を用いたビデオ符号化により、前記第1フレームに対応する遅延ビデオデータの符号化を制御するステップとを有することを特徴とするビデオエンコーダの制御方法。

【請求項24】 前記フィールド差の絶対値の和が所定しきい値を超えたことから複数のフィールドの間のシー

ン変化を検出するステップと、
前記シーン変化の検出にตอบสนองしてシーン変化フラグ信号を生成するステップと、
前記シーン変化フラグ信号を用いて、前記遅延ビデオデータの符号化を制御するステップを実行するステップとをさらに有することを特徴とする請求項 2 3 に記載の方法。

【請求項 2 5】 前記方法は、和変化状態から解像度選択フラグ信号を生成するステップをさらに有し、
前記処理ステップは、
同一パリティを有する連続するフィールドのピクセル間のフィールド差の絶対値の和を計算するステップと、
前記同一パリティを有する連続するフィールドのピクセル間のフィールド差の絶対値の和の変化を和変化状態として検出して解像度選択フラグ信号を生成するステップとを有し、
前記制御するステップは、
前記解像度選択フラグ信号を用いて前記第 1 フレームの符号化の解像度を変化させるステップを有することを特徴とする請求項 2 3 に記載の方法。

【請求項 2 6】 前記入力ビデオデータ内にフィルムデータが存在しないことを検出するステップと、
フィルムデータが検出されない場合にフィルムフラグ信号を生成するフィルムフラグ生成ステップと、
前記フィルムフラグ信号を用いて、前記遅延ビデオデータの符号化を制御するステップを実行するステップとをさらに有することを特徴とする請求項 2 3 に記載の方法。

【請求項 2 7】 前記和から関連フィールドのうちの冗長フィールドを検出するステップをさらに有し、
前記フィルムフラグ生成ステップは、前記冗長フィールドに対応してフィルムフラグを生成し、
前記処理ステップは、
前記第 1 フィールドと、前記第 1 フィールドと同じパリティの連続するフィールドとのピクセル間のフィールド差の絶対値の和を計算するステップを有することを特徴とする請求項 2 6 に記載の方法。

【請求項 2 8】 前記和を用いて前記第 1 フレームから次のフレームへのシーン変化を検出するシーン変化検出ステップをさらに有することを特徴とする請求項 2 7 に記載の方法。

【請求項 2 9】 前記処理ステップは、
前記第 1 フィールドと連続するフィールドとのローパスフィルタイメージ値のサブサンプリング値を計算するステップを有し、
前記シーン変化検出ステップは、
前記第 1 フィールドと連続するフィールドとのローパスフィルタイメージ値のサブサンプリング値のフィールド差の絶対値の最大値を用いて局所的に変化する領域を検出するステップと、

所定値を越える前記フィールド差の絶対値の最大値が所定値を超えた場合にフィルムフラグを生成するステップとを有することを特徴とする請求項 2 8 に記載の方法。

【請求項 3 0】 前記処理ステップは、
各フィールドの平均ピクセル値を計算するステップを有し、
前記シーン変化検出ステップは、
前記第 1 フィールドと連続するフィールドとの平均ピクセル値を用いてビデオフェードを判定するステップと、
平均ピクセル値が所定値を超えた場合にビデオフェードフラグ信号を生成するステップとを有することを特徴とする請求項 2 8 に記載の方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】本発明は、ビデオエンコーダに関し、特にビデオ信号中の変化を検出し、符号化を実行するビデオエンコーダを制御する方法とシステムに関する。

【0 0 0 2】

20 【従来の技術】画像圧縮システムでは、最大の画像品質を維持しながら最小のビット数にビデオ信号をデジタル化している。M P E G (Motion Picture Experts Group) 標準は、画像圧縮に有効な技術を規定している。そのある技術では、ビデオエンコーダ内で画像圧縮を提供しつつ、この圧縮時にデータの冗長性を維持している。

【0 0 0 3】1 秒間当たり 2 4 フレームの入力ビデオイメージ画像のフィルム材料を公知のプロセスでもって、1 秒間当たり 6 0 フィールドのビデオ信号に変換する、この公知の技術は、例えば、3 : 2 のプルダウン方式で、これは、各ビデオフレームをビデオの 3 個のフィールドと 2 個のフィールドでもって交互に記録している。しかし、この 3 : 2 のプルダウン方法は、その変換に際しては冗長となる。このような冗長性を検知しそれを除去することは、情報の喪失を招くことなく 5 個のフィールド毎に 1 個のフィールドを取り除くことを意味する。しかしこれにより記録し処理されるべきビデオイメージデータを 2 0 % 低減することができる。テレビ放送材料の 9 0 % は、フィルムソースから得られたものであるために、イメージ情報を喪失することなく、2 0 % ものビデオイメージデータを低減できることは、ビデオ符号化に対し、データ容量を大きく削減する利点がある。

【0 0 0 4】しかし、イメージフィールド内のこのような冗長性を検知し、それを取り除くために、フィルムとビデオの混合物あるいは異なるフィルムの組み合わせにおける 3 : 2 のプルダウンの誤った検知は、例えばイメージ復号化のような後続の圧縮イメージの処理に際し、許容できないほどのアーティファクト（偽物）が再構成されることになる。

【0 0 0 5】来入ビデオ情報の変化の予測と符号化パラメータのその対応する修正の予測は、従来公知のもので

ある。このような予測方法では、フレーム内 (intra frames) すなわちリフレッシュフレームの誤った配置、あるいは I フレームの誤った配置により、後続の復号化に際しその品質を劣化させることになる。例えば通常のビデオフレームは、少なくともその前のフレームから通常予測できるものであり、そしてビデオイメージを受信機が得ることを容易にするために、受信機には周期的に送信される内部フレーム (intra frames) が付属するものである。このようなフレーム内符号化およびシーンの変化からのフレームの符号化には、通常のビデオフレームよりは多くのビットを必要とする。その理由は、このようなシーンが変化するフレームは、ほとんど予測できないフレームとなるからである。フレーム内符号化が、シーンの変化の直前あるいは直後に発生する場合には、符号化に必要な平均ビットレートは、符号化の品質を落とすレベルまで増加し、その結果、復号化に際し、目障りなアーティファクト (偽物) が発生することになる。

【0006】

【発明が解決しようとする課題】従って、本発明の目的は、特にビデオ信号中の変化を検出し、フレーム信号の効率的な符号化方法を提供することである。

【0007】

【課題を解決するための手段】本発明のビデオエンコーダを制御するシステムのプロセッサは、遅延ビデオデータを生成するために、所定数 N ($N > 1$) だけ入力ビデオデータを遅延させる多フィールド遅延回路と、第 1 フレームと後続のフレームの統計値を生成し、この統計値から制御信号を生成するために前記の入力ビデオデータを処理する統計値生成機と、エンコーダモジュールと、前記の制御信号にตอบสนองして前記エンコーダモジュールにより第 1 フレームに相当する遅延ビデオデータの符号化を制御するレートコントローラとからなる。前記の統計値生成機は、フィールドピクセル間の差の絶対値の和を計算し、サブサンプルされたローパスフィルタイメージ値を計算する。そしてプルダウンディテクタが、それに組み込まれ前記の和を用いて、関連フィールド内の冗長フィールドを検知し、冗長フィールドに対応する制御信号として冗長性フラッグを生成する。

【0008】シーン変化ディテクタが、それに組み込まれ前記の和を用いて、第 1 フレームと後続のフレームからのシーンの変化を検知し、制御信号としてこのシーンの変化フラッグを生成する。統計値生成機は、各フレームの平均ピクセル値を決定し、フェードディテクタは、ビデオフェードを決定するために第 1 フィールドと後続のフィールドの平均ピクセル値を用いる。

【0009】解像度セクタは、前記の和を用いて解像度選択信号を生成し、これをレートコントローラが用いて後続のフレームの符号化の解像度を決定する。プロセッサからの出力は、フィルムフラッグ、シーン変化フラッグ、フェードフラッグを含み遅延ビデオデータの符号

化を制御するためにレートコントローラに送られる。

【0010】また本発明は、複数のフレームに対応する入力ビデオデータを符号化するビデオエンコーダを制御する方法を開示する。本発明の方法は、フレームの所定数 N ($N > 1$) だけ入力ビデオデータを遅延させて遅延ビデオデータを生成するステップと、第 1 ビデオフレームの統計値を生成するために入力ビデオデータを処理するステップと、前記統計値を用いてビデオエンコーダにより第 1 フレームに相当する遅延ビデオデータの符号化を制御するステップを含む。

【0011】

【発明の実施の形態】本発明を説明する前に、一般的事項をまず説明する。ピクセル (ペル) のブロックは、イメージ (画像) を構成するために行列に並んでいる。そして各ピクセルは、3つの要素に関連している。すなわち輝度 Y と赤色差 C_r と青色差 C_b である。フィールドに配置されたビデオデータは、ナショナルテレビジョンシステム委員会「National Television System Committee (NTSC)」標準のビデオでは、約 59.94 Hz (約 60 Hz) で、フェイズアルタネーションライン「Phase Alternation Line (PAL)」標準のビデオでは 50 Hz で動作する。この NTSC 標準と PAL 標準の基では、フィールド対がフレームを構成する。FIELD 1 は、表示されるべき第 1 フィールド (奇数フィールド) と称し、FIELD 2 は、表示されるべき第 2 フィールド (偶数フィールド) と称する。かくして各フィールドは、パリティ、すなわち偶数あるいは、奇数に関係している。

【0012】このようなフィールドは、また内部フィールド (intra field) (1 フィールド) と、予測フィールド (P フィールド) と双方向フィールド (B フィールド) としてカテゴリに分けられ、これらは、1 フレーム、P フレーム、B フレームと同様に定義される。

【0013】図 1 において、エンコーダ 10 は、入力装置からのコマンドあるいは他の入力を受信するエンコーダ制御モジュール 11 と、ビデオ入力内のフィルムフレームおよびシーン変化を検知するプリプロセッサ 12 と、レートコントローラ 14 と、エンコーダモジュール 16 と、予測モジュール 18 と、フォーマッタ 20 と、知覚モデルモジュール 22 と、動き予測モジュール 24 と、デコーダモジュール 26 とを有する。エンコーダ 10 のこれらの構成部品は、従来公知のもので例えば米国特許第 5, 144, 423 号、第 5, 231, 484 号、第 5, 247, 363 号、第 5, 293, 229 号、第 5, 325, 125 号等を開示されている。

【0014】ビデオエンコーダ制御システムおよびここに開示した方法においては、プリプロセッサ 12 は、エンコーダ制御モジュール 11 により処理された入力ビデオデータとコマンド入力とを受信し、フィルムソースに対応するビデオデータからの冗長フィールドを取り除

10

20

30

40

50

く。レートコントローラ14は、プリプロセッサ12からのフラッグのようなデータ信号を受信し、符号化を実行するためにエンコーダ10の動作を制御する。レートコントローラ14は、また動作バンド幅内に符号化されたビットレートを維持するために、外部システムとエンコーダ10との通信を制御する。エンコーダモジュール16は、プリプロセッサ12からのビデオデータと予測モジュール18からの予測値を受信し、前処理されたビデオデータを符号化する。フォーマッタ20は、ビデオフレームのピクセルのブロックと様々なデータフィールドとを組み合わせることで出力チャンネルを介して出力できるように符号化出力信号を生成する。

【0015】知覚モデルモジュール22は、符号化プロセスのために符号化パラメータを計算し、動き予測モジュール24は、現行のピクセルのブロック内のビデオデータと、先行のイメージデータとをブロックマッチングさせて動きファクタを生成する。デコーダモジュール26は、この符号化プロセスから再構成された予測エラーを生成して、復号化イメージを生成する。

【0016】図2に示した実施例においては、プリプロセッサ12は、ルックアップテーブルモジュール28を有し、このルックアップテーブルモジュール28は、CCIR-601の標準フォーマットでもって入力ビデオデータを受信して、公知の任意のガンマ補正、ペダスタル調整、コントラスト強調等の動作を実行する。様々なテーブルが、ルマ(luma)信号と彩度(chroma)信号のためにルックアップテーブルモジュール28内に納められている。ルックアップテーブルモジュール28を介して垂直クロッピングモジュール30に転送される入力ビデオ信号は、入力ビデオデータをクロップ(生成)する。

【0017】例えば、NTSC標準で入力ビデオデータを処理するためには、この入力ビデオデータは、480本にクロップ(生成)される。一方、PAL標準で入力ビデオデータを処理するためにはこの入力ビデオデータは、576本にクロップされる。ここでは全て576本の活性のラインが用いられる。垂直クロッピングモジュール30は、高精細テレビ標準「high definition television standards (HDTV)」, EGA, VGA, スーパーVGA等の特定のビデオ標準の要件に従って入力ビデオデータをクロップする。これらの標準は、公知のものである。例えば、MPEG標準は、MPEGテストモジュール4の「Coded Representation of Picture and Audio Information」, ISO-IEC/JTC1/SC29/EG11, CCITT SG XV, Working Party XV/1, Document AVC-445b, February 1993.」に開示されている。

【0018】NTSCソースからのビデオデータに対しては、垂直方向にクロップされたビデオデータは、3:2プルダウンプロセッサ32が受信し、3:2のプルダウンを検知し、冗長フィールドのような入力データフィ

ールドを再順序化し、取り除くために処理し、そして1秒当たり24フレームのビデオデータを生成する。この3:2プルダウンプロセッサ32は、3:2のプルダウン、再順序化とフィールドの除去を実行する、これに関しては、米国特許第5,317,398号に開示されている。実施例においては、この3:2プルダウンプロセッサ32は、遅延ビデオデータを生成する。そして、この3:2プルダウンプロセッサ32の他の機能は図3でもって説明する。

【0019】図2において、3:2プルダウンプロセッサ32からの遅延ビデオデータは、垂直フィルタ/サブサンブラ34で受信し彩度サブサンプリング(chroma subsampling)を実行する。本明細書においては、サブサンプリングは、係数1のサンプリングを含むものとする。

【0020】エンコーダ10のエンコーダモジュール16は、公知の方法により4:2:0のビデオデータを符号化する。これに関しては、米国特許第5,253,056号と第5,270,813号を参照のこと。この彩度チャンネル解像度は、水平方向および垂直方向の両方においてルマ解像度の半分である。4:2:2のフォーマットの入力ビデオデータに対しては、垂直フィルタ/サブサンブラ34がエンコーダ10により使用される彩度チャンネルを処理する。

【0021】この遅延ビデオデータがフィルムあるいはノンフィルムに対応するときには、3:2プルダウンプロセッサ32は、フィルム状態あるいはノンフィルム状態をフィルムフラッグによって示す。垂直フィルタ/サブサンブラ34は、このフィルムフラッグにตอบสนองして、遅延フィルムデータを累進的フィルムとしてのビデオデータが所定の4個のタップフィルタを用いてフレームベースでもって彩度フィルタ処理されるように遅延フィルムデータを処理する。このビデオデータが飛び越された(interlaced)ビデオデータである(したがって累進的フィルムデータではない)場合には、このビデオデータは、各フィールドに対し独立に垂直方向の彩度フィルタ処理でもって完全な時間レートでもって符号化される。この実施例においては、奇数フィールドは、所定の7タップフィルタでもってフィルタ処理され、偶数フィールドは、所定の4タップフィルタでもってフィルタ処理され、この所定のタップフィルタは対称的なものである。

【0022】垂直フィルタ/サブサンブラ34からのこの彩度フィルタ処理されたビデオデータは、適応形プレフィルタ36と水平フィルタ/サブサンブラ38とその後水平クロッピングモジュール40により処理されて前処理されたビデオデータを生成し、そしてこれがエンコーダモジュール16に出力される。適応形プレフィルタ36と水平フィルタ/サブサンブラ38と水平クロッピングモジュール40の動作は図5でもって説明する。

【0023】次に図3において、3:2プルダウンプロ

10

20

30

40

50

セッサ32は、垂直方向にクロップされた入力ビデオデータ内の冗長フィールドを検知する。3:2のプルダウンは1秒あたり24フレームで記録されたフィルムを約60Hzで動作するNTSCテレビシステムに表示する際に用いられ、これは、20分の1秒あるいは30分の1秒の間のいずれかでフィルムの交互のフレームを表示することにより達成される。そしてこのNTSCテレビカメラは、各フィルムフレームに対しそれぞれ3個のフィールドあるいは2個のフィールドのいずれかを公知の方法で記録する。

【0024】この3:2のプルダウンプロセスにおいては、冗長フィールドが生成される。この3:2プルダウンプロセス32は、入力ビデオデータを受信し、所定数のフィールドN(N>1)だけ入力ビデオデータを遅延させることにより遅延ビデオデータを生成する。第1フィールド遅延42と第2フィールド遅延44は、それぞれこの入力ビデオデータを1フィールドだけ遅延させて、冗長な2フィールド遅延したビデオデータを生成する。その後、この2フィールド遅延したビデオデータは、8フィールド遅延46により遅延されて、遅延ビデオデータを生成し、これは全部で10フィールドすなわち5フレームだけ遅延された入力ビデオデータとなる。

【0025】統計値生成機48は、入力データと2フィールドだけ遅延したデータを第2フィールド遅延44から受信してそこから統計値を生成する。

【0026】このフレーム間統計値と、遅延ビデオデータ（この実施例では、5個のフレーム遅延した）によりエンコーダモジュール16が統計値を処理し、遅延したビデオデータの後続の符号化を調整できるようにする。この統計値を用いてフラッグおよび他のデータ信号のような制御信号でもって来入ビデオデータをカテゴリに分け、エンコーダ10のルックahead機能を提供することにより、符号化プロセス内のどのような圧縮も強化できる。エンコーダ10はこのような統計値、フラッグ等を用いるに際し、入力ビデオ信号内の冗長性を除去し、ビデオイメージ内のシーンの変化のようなビデオイメージの比較的困難な部分を符号化する前に予測し禁止行動を取ることができる。フィルム、あるいはシーン変化を表わすようなプリプロセス12からのフラッグあるいは他の指示に応じてレートコントローラ14の指示からエンコーダ10は、次の内部フレームを再度スケジュール化したり整合することができ、そのため符号化イメージ品質を改善できる。

【0027】この実施例においては、入力ビデオデータは、Y、C_b、C_rを有するビデオピクセルのフィールドに対応し、そこから統計値生成機48は、各入力ビデオフィールドに対し、次の統計値を計算する。

【0028】1) 同一のパリティ（偶数、奇数）の連続する入力フィールドのピクセル間の差の絶対値の和、

【0029】2) 同一パリティの連続する入力フィール

ドのローパスフィルタ処理（および可能ならばサブサンプル化された）されたイメージ値のピクセル間の差の最大値。この実施例においては、入力フィールドのピクセルのブロックの値のブロック平均が計算され入力フィールドのサブサンプルされたローパスフィルタ処理されたイメージのサンプルとして用いられる。

【0030】3) 入力フィールドにおけるピクセルの平均値。

【0031】この統計値生成機48は、デジタル信号処理(DSP)回路あるいはチップのようなデジタル信号処理手段を有し、これはAT&T社から市販されているDSP1610チップを用いることが可能である。更に公知のデジタル信号処理ソフトウェアあるいは他の等価のデジタル信号処理手段も統計値生成機48で使用する上記の統計値を決定する。そしてこれはプルダウンディテクタ50とシーン変化ディテクタ52とフェードディテクタ53により使用される。

【0032】図4において、統計値生成機48は、入力データと2フィールド遅延化入力データを受信する。この統計値生成機48は、差の絶対値計算機54を有し入力データと2フィールド遅延化入力データとの差の絶対値を生成し、フィールドの和計算機56がこの絶対値を用いて、フィールドに亘る絶対値の和を生成する。その後この和は、プルダウンディテクタ50とシーン変化ディテクタ52に出力される。

【0033】この統計値生成機48は、ローパスフィルタ/サブサンプル58と60とを有し、それぞれ入力データと2フィールド遅延化入力データとを受信する。これらから生成されたサブサンプル化されたローパスフィルタ処理されたイメージ値が差の絶対値計算機62に出力されてローパスフィルタ処理されたイメージ値間の差の絶対値を生成する。フィールドの最大値検出機64は、フィールドに亘って最大絶対値を決定し、この最大絶対値がプルダウンディテクタ50に出力される。

【0034】ローパスフィルタ/サブサンプル60は、またフィールドの平均値計算機66を有し、このフィールドの平均値計算機66は、フィールドに亘って入力データの平均値を決定し、この平均値がフェードディテクタ53に出力される。

【0035】本発明のビデオエンコーダ制御システムの以下のシステムとその方法と以下の実施例は、入力ビデオデータから生成された統計値を用いてフィルムデータシーン変化ビデオフェージング等を検出して符号化解像度に適応させようとするものである。

【0036】フィルム検知

第1実施例において、3:2プルダウンプロセス32のプルダウンディテクタ50は、入力ビデオデータ中に存在するフィルムデータを入力ビデオデータの交互のフィールド間のピクセル間の値の絶対値の和を用いて検出する。静止フィールドあるいは冗長フィールド(still

or redundant field) のビデオデータにより、統計値生成機48は、絶対値の差の低い和を生成し、一方静止イメージ(stationary image) のビデオデータにより統計値生成機48は、差の絶対値の和(0に近づく)を生成する。更にこのビデオデータが3:2の変換により変換フィルムイメージから生成された場合には、各5フィールド毎に、5番目のフィールドからの和は小さく、一方5フィールドの残りの4フィールドからの和は大きい。

【0037】プルダウンディテクタ50は、ビデオデータ内のフィルムの存在をそこから生成された和の相対値を比較することにより検出して、第1所定値により差を出しこれによりプルダウンディテクタ50がレートコントローラ14に送信されるフィルムフラッグを生成するようにする。

【0038】プルダウンディテクタ50は、統計値生成機により計算されたローパスフィルタ処理されたイメージの差の絶対値の和とその最大絶対値を用いて冗長なビデオフィールドを検出する。冗長なフィールドの位置の検出に基づいて、10個のフィールドのグループがフィルムあるいはノンフィルムとして区別される。

【0039】一対の入力フィールド間の差の絶対値の和は、マクロスケールで一対のフィールド間の相対的ミスマッチを表わすことになる。この差の絶対値の和が小さい値の時は、一対のフィールドに冗長フィールドが存在する場合だけである。

【0040】ローパスフィルタ処理されたイメージの差の最大絶対値は、この一対の入力フィールドのローカル領域内の相対的ミスマッチを表わす。この差の最大絶対値が小さい場合は、一対のフィールド内に冗長フィールドが存在する場合のみである。

【0041】この入力ビデオデータが3:2のプルダウンにより変換されたフィルムイメージから生成された場合には、10個のフィールドのグループの内、5番目のフィールドと10番目のフィールドが冗長フィールドである。更に、6番目のフィールドと8番目のフィールドとの間のミスマッチの程度は、7番目のフィールドと9番目のフィールドの間のミスマッチの程度によく似ている。4番目のフィールドと6番目のフィールドとの間のミスマッチの程度は、5番目のフィールドと7番目のフィールドとの間のミスマッチの程度と似ている。第1フィールドと第3フィールドとの間のミスマッチの程度は、第2フィールドと第4フィールドとの間のミスマッチの程度に似ている。

【0042】この実施例においては、プルダウンディテクタ50は、8ユニットの長さの内部統計の先入れ先出し(FIFO)の待ち行列を6個有しており、そして10個の最も最近の入力フィールドの統計値を記憶している。その待ち行列は次の通りである。

【0043】a) $\{D_r[0], \dots, D_r[7]\}$

は、ルマ信号(luma signal)のフィールド差の絶対値

の和を記憶する。

【0044】b) $\{D_{cr}[0], \dots, D_{cr}[7]\}$

は、C_r彩度信号(chroma signal)のフィールド差の絶対値の和を記憶する。

【0045】c) $\{D_{cb}[0], \dots, D_{cb}[7]\}$

は、C_b信号のフィールド差の絶対値の和を記憶する。

【0046】d) $\{d_r[0], \dots, d_r[7]\}$ は、ローパスフィルタ処理されたルマ信号のフィールド差の最大絶対値を記憶する。

【0047】e) $\{d_{cr}[0], \dots, d_{cr}[7]\}$

は、ローパスフィルタ処理されたC_r彩度信号のフィールド差の最大絶対値を記憶する。

【0048】f) $\{d_{cb}[0], \dots, d_{cb}[7]\}$

は、ローパスフィルタ処理されたC_b彩度信号のフィールド差の最大絶対値を記憶する。

【0049】上記の説明において、所定のフィールドに対するフィールド差の絶対値は、フィールド全体に亘って計算される。

【0050】プルダウンディテクタ50は、また状態変数 Ψ を保持し、そして状態変数 Ψ は、8フィールド遅延46から最も最近出力されたフィールドのフィルムモードを表わす。この状態変数 Ψ は、最も最近処理されたフィールドがノンフィルムの場合には、0として定義され状態変数 Ψ は、1から10の範囲の1つの値を取り、10フィールドの3:2プルダウンパターン内の出力フィールドの順序を表わす。最初に状態変数 Ψ は、0にセットされている。

【0051】各入力フィールドに対しプルダウンディテクタ50は、次の統計値を用いる。

【0052】1) D_r' , フィールドについての入力輝度データと2フィールド遅延化輝度データとの間のピクセル値のフィールド差の絶対値の和

【0053】2) D_{cr}' , フィールドに亘って入力C_r彩度データと2フィールド遅延化C_r彩度データとの間のピクセル値のフィールド差の絶対値の和

【0054】3) D_{cb}' , フィールドに亘って入力C_b彩度データと2フィールド遅延化C_b彩度データとの間のピクセル値のフィールド差の絶対値の和

【0055】4) d_r' , フィールドに亘って入力輝度データと2フィールド遅延化輝度データのローパスフィルタ処理されたイメージのピクセル値の間のフィールド差の絶対値の最大値

【0056】5) d_{cr}' , フィールドに亘って入力C_r彩度データと2フィールド遅延化C_r彩度データのローパスフィルタ処理されたイメージのピクセル値の間のフィールド差の絶対値の最大値

【0057】6) d_{cb}' , フィールドに亘って入力C_b彩度データと2フィールド遅延化C_b彩度データのローパスフィルタ処理されたイメージのピクセル値の間のフィールド差の絶対値の最大値

10

20

30

40

50

【0058】このプルダウンディテクタ50は、統計値のFIFOを次式に従って更新する。

$$D_Y[n] = D_Y[n-1]$$

$$D_{Cr}[n] = D_{Cr}[n-1]$$

$$D_{Cb}[n] = D_{Cb}[n-1]$$

$$d_Y[n] = d_Y[n-1]$$

$$d_{Cr}[n] = d_{Cr}[n-1]$$

$$d_{Cb}[n] = d_{Cb}[n-1]$$

$n = 1, 2, \dots, 7$ に対し、そして

$$D_Y[0] = D_Y'$$

$$D_{Cr}[0] = D_{Cr}'$$

$$D_{Cb}[0] = D_{Cb}'$$

$$d_Y[0] = d_Y'$$

$$d_{Cr}[0] = d_{Cr}'$$

$$d_{Cb}[0] = d_{Cb}'$$

【0059】現行の入力フィールドが偶数（最底）のフィールドで状態変数 Ψ の現行状態が0あるいは10の時には、プルダウンディテクタ50は、以下のようにフィルム検出を行う、あるいはプルダウンディテクタ50は、フィルム検出を行わずに状態変数 Ψ を1だけ増加させる。

【0060】プルダウンディテクタ50は、式（1）－（4）の状態の全てが満たされた場合のみ8フィールド遅延46から出力された次の10個のフィールドは、フィルムであると決定する。

【0061】

*

表 1

フィルムモード数値	記 述
0	フィールド1／フィールド2
1	フィールド2／フィールド1
2	フィールド1／フィールド2／フィールド1の繰返し
3	フィールド2／フィールド1／フィールド2の繰返し

【0068】シーンの変化の検出

第2の実施例において、統計値生成機48は、情報をプルダウンディテクタ50に送り、このプルダウンディテクタ50は、入力ビデオ内のシーン、すなわちカットの瞬時の変化の発生を検出する。現行のフィールドと2フィールド遅延化フィールド内のピクセル間の差の絶対値の和を用いて、フィールド間の変化の量を分類する。通常のシーンにおいては、このパラメータは、小さく予測不能な動きのシーンの場合でさえも数フレームに亘ってゆっくりと変化する。そのためシーンは、全てのフィールド差が所定のしきい値 T_{low} 以下の場合には、連続的であると見なされ小量 ΔT_{low} だけ連続的な値の間で変化する。

【0069】しかし、複数のシーンの間のカットにおいては、このパラメータは、大きくなる。その理由は、フィールド差を計算するのに用いられる2個のフィールドは、異なったシーンで発生するからである。この場合、このフィールド差は、通常所定のしきい値 T_{high} よりも

*【数1】

【0062】ここで、 T_Y 、 T_{Cr} 、 T_{Cb} 、 t_Y 、 t_{Cr}

、 t_{Cb} 、 R_Y 、 R_{Cr} 、 R_{Cb} 、 R_Y' 、 R_{Cr}' 、

R_{Cb}' は、入力データから前以て得た知識により決定された、則ち、予め存在するしきい値である。

【0063】

【数2】

【0064】例えば、これらのしきい値は、ユーザの入力コマンドあるいは、トレーニングデータによって設定可能である。

【0065】上記の式（1）－（4）の全てが満たされた場合、すなわちこのフィルムテストがフィルム状態が肯定的であることを示した場合には、状態変数 Ψ は、1に設定され、8フィールド遅延46から出力された次の10個のフィールドは、フィルムとしてクラス分けされる。それ以外では、状態変数 Ψ は、0にセットされ、8フィールド遅延46から出力された次の2個のフィールドは、ノンフィルム（すなわちビデオ）として分類され上記の検出プロセスは、新たな入力データに対し繰返される。

【0066】このプルダウンディテクタ50は、また2ビットモードの字数値を生成し、これは、0-3（ベース10）を有し、レートコントローラ14に対し表1に示すようなフレームの再構成を示す。

【0067】

大きい。さらにフィールド差は、交互のフィールド間で計算されるために、シーンのカットが発生すると2つの連続するフィールドは、しきい値 T_{high} 以上の大きなフィールド差を有する。

【0070】そのためシーンの変化の検出は、少なくとも2つの高いフィールド差が、低いフィールド差の後に来るか、前に来るかを検出することによって、容易に検出できる。統計値生成機48のルックahead特性（look-ahead nature）により、後続のフィールド差の数の検出が進められ、シーンの検出の信頼性を向上させることになる。さらにシーンの変化は、それが符号化されるためにエンコーダモジュール16に入力する1あるいは複数のフレーム時間の間に検出される。これによりエンコーダ10がシーン変化の前に符号化パラメータを変えることができるようになる、これは例えば、シーンの変化の前のフレームの符号化品質を劣化させることになるが、復号化されたビデオを見ている者にとっては、目に見えるような品質の変化を引き起こさないため、これに

よりエンコーダ10の符号化効率を改善できるものである。

【0071】これらの変化を検出すると、シーン変化ディテクタ52は、シーンの変化フラッグを出力し、これをレートコントローラ14に送信する。例えば、このシーンの変化フラッグは、符号化内セッティングのようなピクチャータイプを表わす。例えば、検出のある確信的レベルを達成するようなシーンの変化に対しては、このフレーム内フラッグが生成され、レートコントローラ14に送られて、フレーム全体を内部符号化(intra code) 10 する。このフレーム内フラッグシーン変化ディテクタ52から交互の再同期化方法として出力される。統計値生成機48は、周期的カウントを生成するカウンタ

(図示せず)を有し、プリプロセッサ12はこの周期的カウントに応答して一定間隔でフレーム内フラッグを生成する。この交互の再同期化方法においては、プリプロセッサ12は、シーンの変化フラッグに応答してフレーム内フラッグを生成し、レートコントローラ14は、このフレーム内フラッグに応答して、内部フレームをシーンの変化に相当する場所の複数のフレーム内に挿入する。この方法においては、シーン変化ディテクタ52は、シーン変化の相対的位置に応じて、内部フレームの周波数を変化するように選択し、そして、内部フレームを生成し、シーン変化を一致させる。そしてプリプロセッサ12は、このシーン変化フラッグに応答して、カウンタのカウントを修正し、カウントの修正時にフレーム内フラッグを生成する。

【0072】別法として、シーン変化ディテクタ52は、8ビット値をレートコントローラ14に出力してシーン変化の可能性すなわち強さを表わすようにする。この8ビット値は、また他のシーン変化を部分シーン変化として示す。

【0073】フィルムの検出以外に、3:2プルダウンプロセッサ32は、所定の間隔で内部フレームを挿入して、コマーシャルとIPフレーム符号化を別法として挿入することも可能で、これにより累進的なリフレッシュ動作が可能となる。

【0074】ビデオフェージングの検出

第3の実施例においては、フェードディテクタ53はビデオデータ内のビデオフェージングを決定する。フィルムに追加されたビデオフェードの存在とフィルムからフィルムへの遷移あるいはフィルムからビデオへの遷移のフェードレートによりがフィルムの誤り検出を生成する。このフェードディテクタ53は、フェードの前後のシーンの輝度の大きな変化を表わすものとして、平均的ピクセル値を用いて、統計からビデオの遷移を検出する。このフェードディテクタ53は、平均ピクセル値が所定の値以下であることに応答して、ビデオ内のフェードを表わすフェードフラッグを生成する。

【0075】適応形解像度制御

第4の実施例において、本発明のビデオエンコーダ制御システムは、フレーム統計により測定されたシーンの変化に応じて、符号化の解像度を適合させる。3:2プルダウンプロセッサ32の統計値生成機48は、入力ビデオデータのピクセルのフィールド差の絶対値の和を対応するイメージの複雑さの尺度として生成する。別法として、フィールド間のピクセルの変動も計算される。これらの和は、レートコントローラ14に送信され、このレートコントローラ14は、これらの和が所定値を越えるか否かを決定する。この所定の値以上になると、すなわち符号化されるべき画像の複雑性が高いと、レートコントローラ14は、符号化イメージの解像度を落とす解像度選択フラッグを生成する。

【0076】図5において、入力ビデオデータは、適応形プレフィルタ36により処理されるがこの実施例においては、各フレーム毎1度の割合でプレフィルタ選択信号を用いて、レートコントローラ14により選択された256個のフィルタ値あるいは係数の第1所定バンクを用いて行われる。適応形プレフィルタ36は、また3:2プルダウンプロセッサ32からフィルムフラッグを受領し、この入力ビデオデータが累進的フィルムあるいはインタレースされたビデオのいずれかに相当するかを表わすものである。この適応形プレフィルタ36は、フィルムフラッグを用いてフィルタ値の組の第2の所定のバンクからフィルタ値を選択する。

【0077】このフィルムモードフラッグとプレフィルタ選択信号とは、8×8フィルタ68により処理されこの8×8フィルタ68は、フレームベースで8×8のフィルタ処理を行うために8本の水平スキュンラインタップと8本の垂直タップの組からなる。この8×8フィルタ68のタップは、プログラム可能なものである。例えば、1本おきの行列のタップを0に設定すると、8×8のフィルタは、4×4フィルタとなりこれはフィールドフィルタ処理を実行する為にフレームのピクセルに適應できるものである。この8×8フィルタ68の対応する出力値の計算に関連するものは、フレーム当たり2個のフィールドの1つからの値のみである。

【0078】このフィルムフラッグがオフの時には、8×8フィルタ68は、8×4のフィルタとして動作してフィールドベースでフィルタ処理を行う。そしてフィルムフラッグがオンの時には、8×8フィルタ68は、フレームベースでフィルタ処理を行い、タップに係る条件(すなわちタップの1つおきの行列は、0に設定される状態)は除かれる。かくしてプレフィルタ処理は、ビデオに対しては、フィールドベースで、フィルムに対しては、フレームベースで実行される。

【0079】複数のフィルタすなわちフィルタ値の組あるいは、フィルタ係数の組が各タイプのフィルタ処理、例えばフィールドベースのフィルタ処理とフレームベースのフィルタ処理に対し与えられる。このような複数の

フィルタは、 8×8 フィルタ 6 8 のタップを異なる方法でセットすることにより決定できる。入力ビデオデータがフィルム材料であるかあるいは、ビデオ材料であるかに応じてフィールドベースあるいはフレームベースのフィルタ処理に適合する以外に、 8×8 フィルタ 6 8 は、ビデオイメージの複雑化およびエンコーダ 1 0 の状態に基づいて、ビデオデータを符号化する為の困難度に基づいて、フィールドフィルタの値の組、あるいは、フレームフィルタの値の組のいずれかの 1 つから選択することにより適応形となる。このシーンの複雑さは、同一パリティの連続フィールドのフィールド差の絶対値の和により決定される。

【0080】別法として、 8×8 フィルタ 6 8 は、フィールド差の絶対値の和およびレートコントローラ 1 4 により設定される符号化ビットレートとを用いて遅延ビデオデータをフィルタ処理するために複数のフィルタ値の少なくとも 1 つを選択する。

【0081】このフィルタ値の所定の組のあるものは、負であるので、フィルタ処理におけるオーバフローの状態は、 8×8 フィルタ 6 8 の出力をクランプ 7 0 を用いて、0 から 2 5 6 までの 8 ビット範囲内にクランプあるいはラッチすることにより回避できる。

【0082】各 Y と C_r と C_b の値が、上記の方法で異なる組のフィルタ値でもって独立にフィルタ処理される。

【0083】この実施例においては、適応形プレフィルタ 3 6 は 2 : 1 の水平フィルタ/サブサンブラ 3 8 により処理される 7 2 0 個のピクセルのクランプされたデータを出力する。この 2 : 1 の水平フィルタ/サブサンブラ 3 8 は、レートコントローラ 1 4 からの解像度選択フラッグに応答してサブサンプリングのレートを決定する。MPEG 符号化標準を用いた実施例においては、7 0 4 モードと 3 5 2 モードとがサポートされている。7 0 4 モードが選択された場合には、2 : 1 の水平方向サンブラ 7 2 はディスエイブルされ、1 ライン当たり 7 0 4 個のピクセルを有する処理済みのビデオデータが 2 : 1 の水平方向サンブラ 7 2 に入力されて、そこを変化されずに通過する。3 5 2 モードが選択されると Y、C_r、C_b の値が 2 の係数でもってサブサンプル処理され各ラインの第 1 ピクセルが残り 1 ライン当たり 3 5 2 個のピクセルの解像度が提供される。

【0084】別の実施例においては、水平フィルタ/サブサンブラ 3 8 は、フレームフィルタ処理あるいはフィールドフィルタ処理を実行するのと同様な方法でもってフィルムフラッグを用いて適応形のフィルタ処理を実行する。かくしてこの垂直フィルタ/サブサンブラ 3 4 は、彩度のダウンサンプリングを実行し 4 : 2 : 2 の彩度を 4 : 2 : 0 の彩度に変換し、これは処理されるべきビデオデータがフィルムフラッグにより 3 : 2 プルダウンプロセッサ 3 2 により決定されるフィルム材料である

か、ビデオ材料であるかによって決められる。

【0085】

【発明の効果】本発明の解像度制御方法を上記の適応形プレフィルタ 3 6 と組み合わせて用いることにより解像度のより細かな制御が可能となる。さらに入力ビデオデータの遅延からエンコーダ 1 0 のルックアヘッド機能と解像度を制御するために統計を用いる際にエンコーダ 1 0 は、解像度選択フラッグを用いて解像度を向上させ、および少なくとも 1 つのフレームに割り当てられたビット数を減少し、符号化プロセスのアーティファクト（偽物）の生成を抑制するためにシーンの変化の前に解像度選択フラッグを用いて解像度を増加させ、少なくとも 1 つのフレームに割り当てられたビット数を減少することによりシーンの変化に対応できる。

【0086】図 6 に示したように、本発明の方法は、エンコーダ 1 0 を制御する際にステップ 7 4 でエンコーダ 1 0 の制御の開始をし、ステップ 7 6 で入力ビデオデータを受信し、ステップ 7 8 でビデオデータの前処理をし、ステップ 8 8 で遅延ビデオデータと統計を用いてエンコーダ 1 0 の制御を行う。前処理のステップには、N 個のフレームだけビデオデータを遅延させるステップを含む。ここで N (N > 1) を含み、ステップ 8 0 で遅延ビデオデータを生成する。ステップ 8 2 で入力ビデオデータから統計値を生成し、ステップ 8 4 でシーンの変化フラッグ、フィルムフラッグ、ビデオフェードフラッグ等を含むフラッグを生成するためにこの統計値を処理し、ステップ 8 6 でこの統計値から解像度の設定を決定する。

【図面の簡単な説明】

【図 1】本発明のビデオエンコーダのブロック図

【図 2】プリプロセッサのブロック図

【図 3】図 2 の 3 : 2 のプルダウンプロセッサのブロック図

【図 4】図 2 の 3 : 2 のプルダウンプロセッサの統計値生成機のブロック図

【図 5】図 2 のプリプロセッサの要素を表わすブロック図

【図 6】本発明のビデオエンコーダ制御システムの方法と動作を表わすフローチャート図

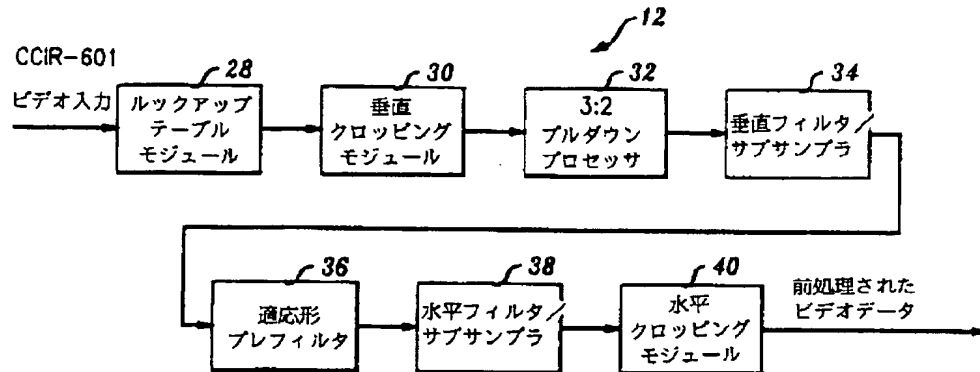
【符号の説明】

- 1 0 エンコーダ
- 1 1 エンコーダ制御モジュール
- 1 2 プリプロセッサ
- 1 4 レートコントローラ
- 1 6 エンコーダモジュール
- 1 8 予測モジュール
- 2 0 フォーマッタ
- 2 2 知覚モデルモジュール
- 2 4 動き予測モジュール
- 2 6 デコーダモジュール

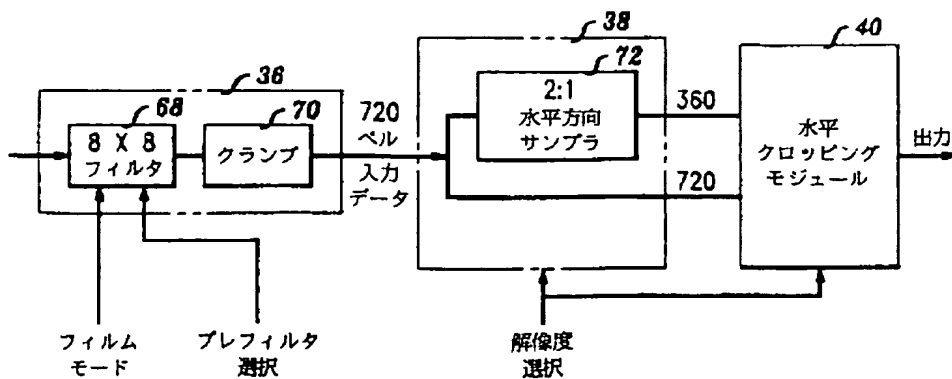
23
 28 ルックアップテーブルモジュール
 30 垂直クロッピングモジュール
 32 3:2 プルダウンプロセッサ
 34 垂直フィルタ/サブサンブラ
 36 適応形プレフィルタ
 38 水平フィルタ/サブサンブラ
 40 水平クロッピングモジュール
 42 第1フィールド遅延
 44 第2フィールド遅延
 46 8フィールド遅延
 48 統計値生成機
 50 プルダウンディテクタ
 52 シーン変化ディテクタ
 53 フェードディテクタ
 54, 62 差の絶対値計算機
 56 フィールドの和計算機
 58, 60 ローパスフィルタ/サブサンブラ

24
 * 64 フィールドの最大値検出機
 66 フィールドの平均値計算機
 68 8×8フィルタ
 70 クランプ
 72 2:1 水平方向サンブラ
 74 ビデオエンコーダの制御開始
 76 入力ビデオデータを受信
 78 ビデオデータの前処理
 80 遅延ビデオデータを生成するためにビデオデータをNフレームだけ遅延させる
 82 入力ビデオデータから統計値を生成する
 84 この統計値を処理してシーン変化フラッグ、フィルムフラッグ、ビデオフェードフラッグを含むフラッグを生成する
 86 この統計値から解像度の設定を決定する
 88 遅延ビデオデータと統計値を用いてエンコーダを制御する

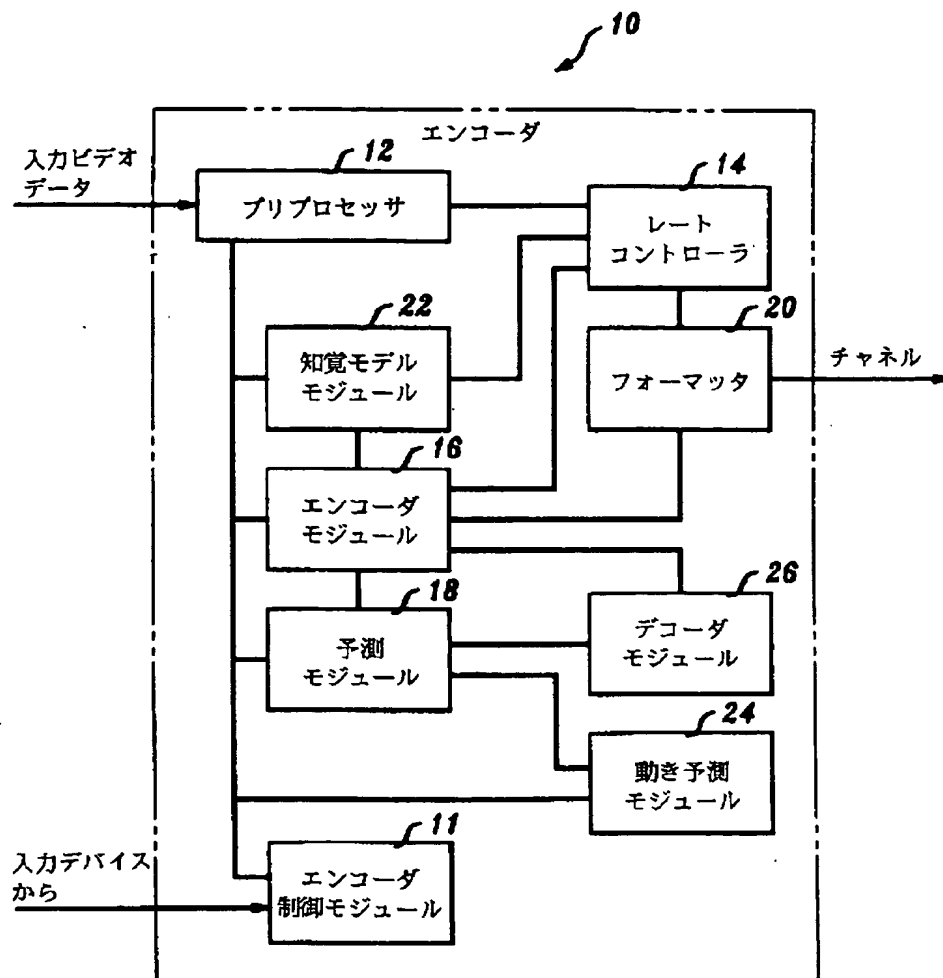
【図2】



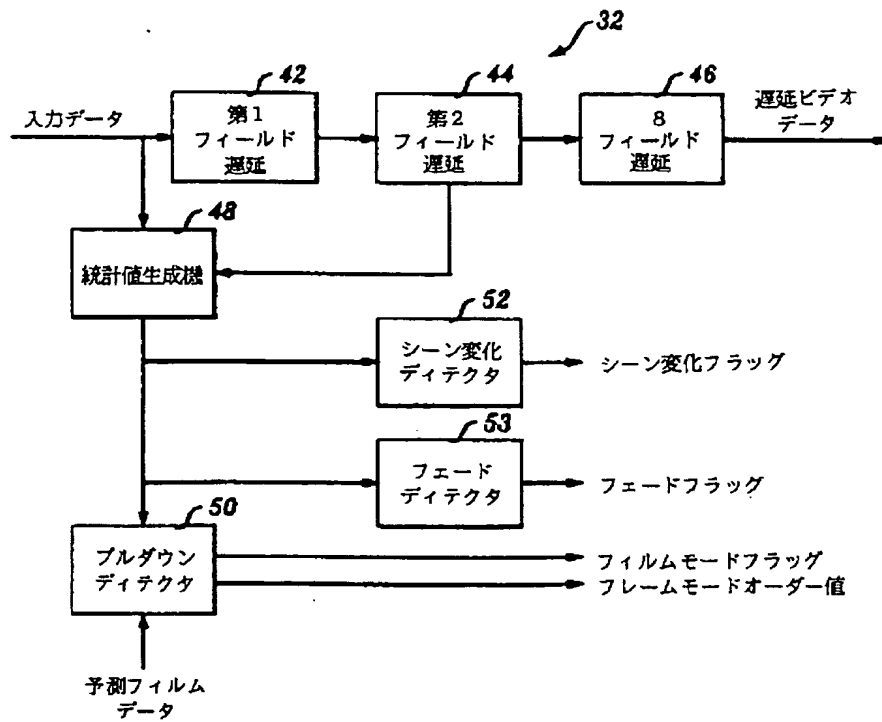
【図5】



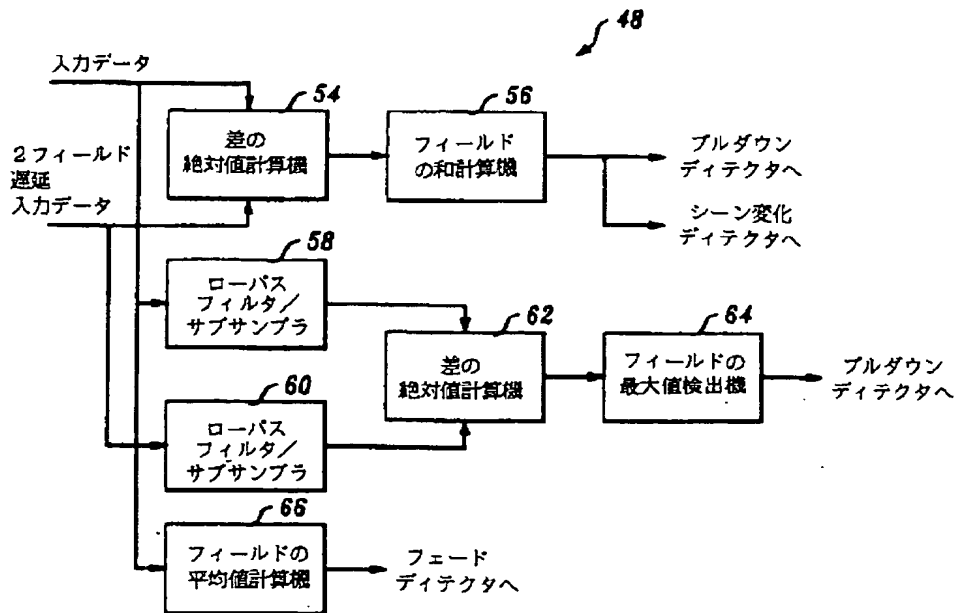
【図1】



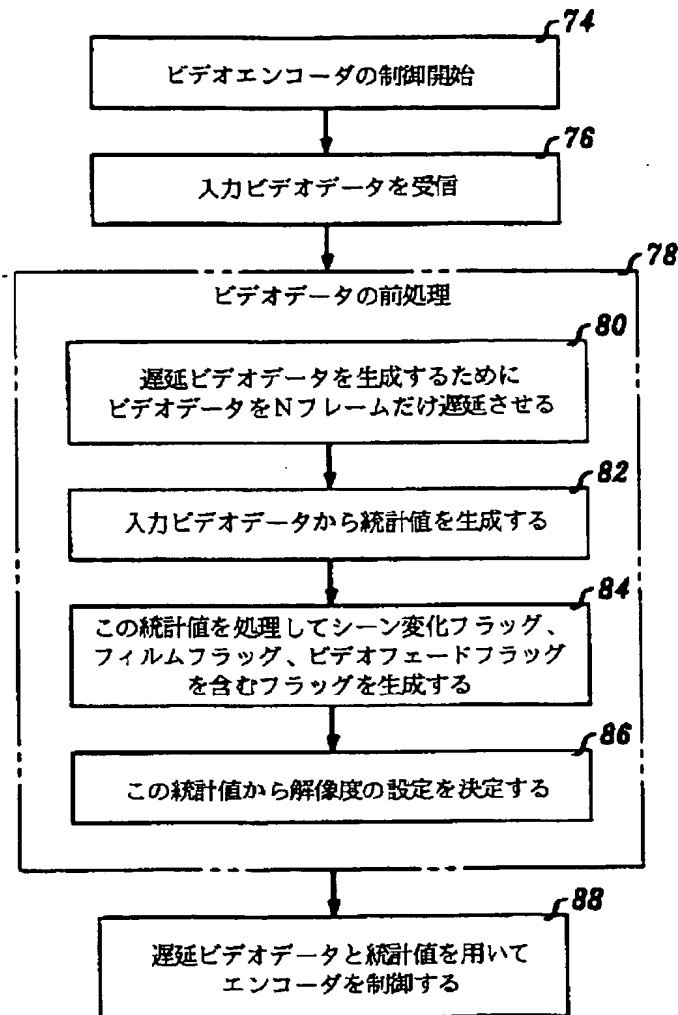
【図3】



【図4】



【図6】



フロントページの続き

(72)発明者 アリレザ ファリッド ファーヤー
 アメリカ合衆国, 07704 ニュージャ
 ー, フェア ハーヴン, ロウリー ド
 ライブ 30

(72)発明者 ラジェッシュ ヒンゴラニ
 アメリカ合衆国, 08550 ニュージャ
 ー, プリンストン ジャンクション,
 インディアン ラン ロード 26

(72)発明者 キム ニジェル マシュー
 アメリカ合衆国, 07060 ニュージャ
 ー, ワッチャン, エッジモント ロー
 ド 54

(72)発明者 デヴィッド トーマス
 アメリカ合衆国, 07901 ニュージャ
 ー, サミット, ラッセル プレイス
 16エー

(72)発明者 シューワイ ウー
 アメリカ合衆国, 08829 ニュージャ
 ー, ハイ ブリッジ, レイク アヴェ
 ニュー 18

- (56)参考文献 特開 平 6 - 197273 (J P, A)
特開 平 6 - 133305 (J P, A)
特開 平 6 - 30398 (J P, A)
特開 平 5 - 328317 (J P, A)
国際公開 94 / 16526 (W O, A 1)

- (58)調査した分野(Int. Cl. 7, D B 名)
H04N 7/24 - 7/68
E P A T (Q U E S T É L)

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

(57) [Claim(s)]

[Claim 1] The video encoder control system which is characterized by providing the following and which has a video encoder. It is the multi-field delay circuit to which this video encoder has a processor, the input video data inputted into the aforementioned video encoder has two or more frames containing the 1st frame, each frame has two or more fields, each field has parity, only the frame of N predetermined numbers delays an input video data, and the aforementioned processor generates a delay video data as $N > 1$. Having the statistic generation machine which processes the aforementioned input video data and generates the statistic of the 1st aforementioned frame, the aforementioned statistic is a subsampling value of a low pass filter image value. It is the 1st detector which the aforementioned statistic generation machine generates the average of a block of the pixel of the 1 field as a subsampling value of the low pass filter image value of this field, the aforementioned processor uses the aforementioned maximum in order to perform film detection further, and detects a non film state including the maximum covering the 1 field of the absolute value of the field difference of the pixel value of the subsampling value of a low pass filter image value.

[Claim 2] The aforementioned processor is a video encoder control system according to claim 1 characterized by operating with the aforementioned video encoder so that the aforementioned input video data may be encoded by the MPEG standard.

[Claim 3] It is the video encoder control system according to claim 1 characterized by to have further the 2nd detector which the aforementioned statistic generation machine calculates the sum of the absolute value of the field difference between the pixels of the continuous field which has the same parity, and the aforementioned processor detects the scene change on the following frame from [aforementioned] the 1st frame, and generates a scene change flag signal as a control signal, and the rate controller which controls an encoder module using the aforementioned scene change flag signal.

[Claim 4] It is the video encoder control system according to claim 3 which the aforementioned statistic generation machine calculates the average pixel value of each field, and is characterized by the 2nd detector of the above determining the video fade using the average pixel value relevant to the 1st field of the 1st aforementioned frame, and the continuous field.

[Claim 5] It is the video encoder control system according to claim 1 which the aforementioned statistic generation machine calculates the sum of the absolute value of the field difference between the pixels of the continuous field which has the same parity, and is characterized by the aforementioned video encoder control system having further the resolution selector to which the resolution of the 1st-frame [aforementioned] coding is changed using the sum of the absolute value of the aforementioned field difference.

[Claim 6] It is the video encoder control system according to claim 5 characterized by for the aforementioned resolution selector answering the aforementioned sum, generating a resolution selection signal, and for the aforementioned processor answering the aforementioned resolution selection signal, and having further the sampler which samples the aforementioned delay video data in the resolution corresponding to this resolution selection signal.

[Claim 7] It is the video encoder control system according to claim 1 which the aforementioned statistic generation machine calculates the sum of the absolute value of the field difference between the pixels of the continuous field which has the same parity, and the 1st detector of the above detects the redundant field in the field related using the aforementioned sum, generates the redundancy flag signal corresponding to this redundant field, and is characterized by for the aforementioned processor to have further the rate controller which controls an encoder module using the aforementioned redundancy flag signal.

[Claim 8] The 1st detector of the above is a video encoder control system according to claim 7 characterized by performing 3:2 pulldown detection to an input video data.

[Claim 9] It is the video encoder control system according to claim 8 characterized by for the 1st detector of the above answering the aforementioned 3:2 pulldown detection, generating the film flag signal which shows whether film data were detected in the aforementioned input video data, and for the aforementioned processor answering the aforementioned film flag signal, and having further the pre-filter which carries out filtering of the aforementioned delay video data using two or more filter values.

[Claim 10] The aforementioned pre-filter is a video encoder control system according to claim 9 characterized by choosing at least one of two or more aforementioned filter values using the sum and the encoder bit rate of an absolute value of a field difference in order to carry out filtering of the aforementioned delay video data.

[Claim 11] The 1st detector of the above is a video encoder control system according to claim 7 characterized by judging the field which changes locally using the maximum of the absolute value of the field difference relevant to the 1st aforementioned frame and a continuous frame.

[Claim 12] It is the video encoder control system according to claim 11 which it has the following, the aforementioned processor answers the aforementioned periodic count and generates an inner flame flag, the aforementioned processor answers the aforementioned scene change flag, and generates an inner flame flag, and the aforementioned rate controller answers the aforementioned inner flame flag, and is characterized by inserting an inner flame in the position corresponding to the scene change in two or more aforementioned frames. The aforementioned processor is the 2nd detector which detects the scene change on the following frame from [aforementioned] the 1st frame, and generates a scene change flag. The counter which generates a periodic count

[Claim 13] The aforementioned pre-filter is a video encoder control system according to claim 12 characterized by having the video-data clamp which clamps the delay video data by which filtering was carried out to predetermined within the limits in order to avoid overflow of a filter.

[Claim 14] The aforementioned processor is a video encoder control system according to claim 13 characterized by answering the aforementioned scene change flag, updating the count of the aforementioned counter, answering renewal of this count, and generating an inner flame flag.

[Claim 15] The video encoder characterized by providing the following. In the video encoder which detects the film in the input video data corresponding to two or more frames which should be encoded, and scene change The input video data inputted into the aforementioned video encoder It has two or more frames containing the 1st frame, each frame has two or more fields, and each field has parity. the aforementioned video encoder The multi-field delay circuit to which only the frame of N predetermined numbers delays an input video data, and generates the 1st delay video data as $N > 1$ It has a low pass filter, and has the statistic generation machine which processes the aforementioned input video data and generates the statistic of the 1st aforementioned frame as a control signal, and the aforementioned statistic is a subsampling value of a low pass filter image value. The sum of the absolute value of the field difference of the subsampling value of a low pass filter image value The maximum covering the 1 field of the absolute value of the field difference of the pixel value of the subsampling value of a low pass filter image value is included. the aforementioned statistic generation machine The sum of the absolute value of the field difference between the pixels of the continuous field which has the same parity is calculated. The aforementioned video encoder generates the film flag which shows whether the non film state of the 1st aforementioned frame was detected from the aforementioned maximum, and film data were further detected using the aforementioned statistic. Since the sum of the aforementioned absolute value exceeded the predetermined threshold, detect two or more aforementioned inter-frame scene change, and a scene change flag is generated. The 1st detector which detects two or more aforementioned inter-frame video fade states, and generates a video fade flag. An encoder module, the generated film flag and the generated scene change flag, and the generated video fade flag are answered. The rate controller which transmits one control signal to the aforementioned encoder module even if this ** cannot be found, in order to control that generate at least one control signal and the aforementioned encoder module encodes the aforementioned 1st delay video data

[Claim 16] It is the video encoder according to claim 15 characterized by having further the resolution selector which the aforementioned video encoder answers the aforementioned sum and generates a resolution selection signal, and for the aforementioned rate controller answering the aforementioned resolution selection signal, and changing the resolution of the 1st-frame [aforementioned] coding.

[Claim 17] It is the video encoder according to claim 15 characterized by for the aforementioned video encoder having further the control module which processes an input command, and for the 1st detector of the above answering the film command predicted, and generating a film flag.

[Claim 18] It is the video encoder according to claim 15 which the aforementioned statistic generation machine calculates the sum of the absolute value of the field difference between the pixels of the continuous field which has the same parity, and is characterized by the 1st detector of the above detecting the scene change on the following frame from [aforementioned] the 1st frame using the aforementioned sum.

[Claim 19] It is the video encoder according to claim 18 which the aforementioned statistic generation machine calculates the average pixel value of each field, and the aforementioned video encoder judges a video fade state using the average pixel value relevant to the 1st field of the 1st aforementioned frame, and the continuous field, and is characterized by having further the 2nd detector which generates a corresponding video fade flag.

[Claim 20] It is the video encoder according to claim 15 characterized by for the aforementioned statistic generation machine calculating the sum of the absolute value of the field difference between the pixels of the continuous field which has the same parity, and for the 1st detector of the above detecting the redundant field of the fields related using the aforementioned sum, and generating a film flag corresponding to this redundant field.

[Claim 21] It is the video encoder according to claim 20 characterized by only for the at least 1 field delaying the aforementioned input video data, for the aforementioned multi-field delay circuit generating the 2nd delay video data, and the aforementioned statistic generation machine generating the sum of the absolute value of a field difference from the aforementioned input video data and the aforementioned 2nd delay video data.

[Claim 22] The 1st detector of the above is a video encoder according to claim 20 characterized by judging the field which changes locally using the maximum of the absolute value of the field difference relevant to the 1st aforementioned frame and a continuous frame.

[Claim 23] The control method of the video encoder characterized by providing the following. It is the step to which the aforementioned input video data has two or more frames containing the 1st frame in the method of controlling

the video encoder which encodes an input video data, each frame has two or more fields, each field has parity, only the frame of N predetermined numbers delays an input video data, and the aforementioned method generates a delay video data as $N > 1$. Having the processing step which generates the data in which the aforementioned input video data is processed and the 1st aforementioned frame is shown, the aforementioned data are the subsampling value of a low pass filter image value. The sum of the absolute value of the field difference of the subsampling value of a low pass filter image value. The maximum covering the 1 field of the absolute value of the field difference of the pixel value of the subsampling value of a low pass filter image value is included. The step which generates the film flag signal which the aforementioned method uses the aforementioned maximum further for film detection when film data are not detected in the aforementioned input video data, and is generated, The step which controls coding of the delay video data corresponding to the 1st aforementioned frame by video coding using the data and the aforementioned film flag signal which were generated

[Claim 24] The method according to claim 23 characterized by to have further the step which detects the scene change between two or more fields from the sum of the absolute value of the aforementioned field difference having exceeded the predetermined threshold, the step which answers detection of the aforementioned scene change and generates a scene change flag signal, and the step which performs the step which controls coding of the aforementioned delay video data using the aforementioned scene change flag signal.

[Claim 25] The step which has the step which detects change of the sum of the absolute value of the field difference between the pixels of the continuous field characterized by to provide the following as a sum change state, and generates a resolution selection flag signal, and carries out the aforementioned control is the method according to claim 23 characterized by to have the step to which the resolution of the 1st-frame [aforementioned] coding is changed using the aforementioned resolution selection flag signal. It is the step to which the aforementioned method has further the step which generates a resolution selection flag signal from a sum change state, and the aforementioned processing step calculates the sum of the absolute value of the field difference between the pixels of the continuous field which has the same parity. The same aforementioned parity

[Claim 26] The method according to claim 23 characterized by having further the step which detects that film data do not exist in the aforementioned input video data, the film flag generation step which generates a film flag signal when film data are not detected, and the step which performs the step which controls coding of the aforementioned delay video data using the aforementioned film flag signal.

[Claim 27] It is the method according to claim 26 which it has further the step which detects the redundant field of the related fields from the aforementioned sum, and the aforementioned film flag generation step generates a film flag corresponding to the aforementioned redundant field, and is characterized by for the aforementioned processing step to have the step which calculates the sum of the absolute value of the field difference between the pixels of the 1st field of the above, and the 1st field of the above and the field where the same parity continues.

[Claim 28] The method according to claim 27 characterized by having further the scene change detection step which detects the scene change on the following frame from [aforementioned] the 1st frame using the aforementioned sum.

[Claim 29] The method according to claim 28 characterized by providing the following. It is the step which the aforementioned processing step has the step which calculates the subsampling value of the low pass filter image value of the 1st field of the above, and the continuous field, and detects the field where the aforementioned scene change detection step changes locally using the maximum of the absolute value of the field difference of the subsampling value of the low pass filter image value of the 1st field of the above, and the continuous field. The step which generates a film flag when the maximum of the absolute value of the aforementioned field difference exceeding a predetermined value exceeds a predetermined value

[Claim 30] It is the method according to claim 28 which the aforementioned processing step has the step which calculates the average pixel value of each field, and is characterized by the aforementioned scene change detection step having the step which judges the video fade using the average pixel value of the 1st field of the above, and the continuous field, and the step which generates a video fade flag signal when an average pixel value exceeds a predetermined value.

[Translation done.]

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[The technical field to which invention belongs] About a video encoder, especially this invention detects the change in a video signal, and relates to the method and system which control the video encoder which performs coding.

[0002]

[Description of the Prior Art] In the picture compression system, the video signal is digitized to the minimum number of bits, maintaining the greatest picture quality. The MPEG (Motion Picture Experts Group) standard has specified technology effective in picture compression. With the technology of a certain, the redundancy of data is maintained at the time of this compression, offering picture compression within a video encoder.

[0003] This well-known technology that has the film material of the input video image picture of 24 per for 1 second in a well-known process, and is changed into the video signal of the 60 fields per for 1 second is the pulldown method of 3:2, and this has each video frame in the three fields of video, and the two fields, and is recording it by turns. However, this pulldown method of 3:2 becomes redundant on the occasion of the conversion. Detecting such redundancy and removing it means removing the one field for every five fields, without causing informational loss. However, the video image data which records by this and should be processed can be reduced 20%. That no less than 20% of video image data can be reduced has the advantage which cuts down data capacity greatly to video coding, without losing image information, since 90% of television broadcasting material is obtained from the film source.

[0004] However, in order to detect such redundancy in the image field and to remove it, an artifact (imitation) to the extent that the detection which PURUDAUN of 3:2 in the mixture of a film and video or a different combination of a film mistook is nonpermissible on the occasion of processing of the compression image of consecutiveness like for example, an image decryption will be reconfigured.

[0005] It comes and prediction and prediction of corresponding correction of a coding parameter of change of ON video information are conventionally well-known. The quality is made to deteriorate in a consecutive decryption by such prediction method in the case by the arrangement which the inside of a frame (intra frames), i.e., a refreshment frame, mistook, or mistaken arrangement of I frames. For example, the usual video frame can usually be predicted from the frame in front of it at least, and in order to make it easy that a receiver acquires a video image, the inner frame (intra frames) transmitted periodically is attached to a receiver. Rather than the usual video frame, many bits are needed for coding of the frame from such coding in a frame, and change of a scene. The frame from which, as for the reason, such a scene changes is because it becomes the frame which can hardly be predicted. When generating in immediately after [just before coding in a frame changes / of a scene], an average bit rate required for coding will increase to the level to which the quality of coding is lowered, consequently an artifact (imitation) offensive to the eye will generate it on the occasion of a decryption.

[0006]

[Problem(s) to be Solved by the Invention] Therefore, especially the purpose of this invention is detecting the change in a video signal and offering the efficient coding method of a frame signal.

[0007]

[Means for Solving the Problem] The processor of the system which controls the video encoder of this invention has a multi-field delay circuit from which only a predetermined number N ($N > 1$) delays an input video data in order to generate a delay video data. The statistic generation machine which processes the aforementioned input video data in order to generate the statistic of the 1st frame and a consecutive frame and to generate a control signal from this statistic. It consists of an encoder module and a rate controller which controls coding of the delay video data which answers the aforementioned control signal and is equivalent to the 1st frame with the aforementioned encoder module. The aforementioned statistic generation machine calculates the sum of the absolute value of the difference between field pixels, and calculates the low pass filter image value by which the sub sample was carried out. And it is included in it, and using the aforementioned sum, a pulldown detector detects the redundant field in the related field, and generates a redundancy flag as a control signal corresponding to the redundant field.

[0008] It is included in it, and using the aforementioned sum, a scene change detector detects change of the scene from the 1st frame and a consecutive frame, and generates the change flag of this scene as a control signal. A statistic generation machine determines the average pixel value of each frame, and in order to determine the video fade, the average pixel value of the 1st field and the consecutive field is used for a fade detector.

[0009] A resolution selection signal is generated using the aforementioned sum, a rate controller uses this, and a

resolution selector determines the resolution of coding of a consecutive frame. The output from a processor is sent to a rate controller, in order to control coding of a delay video data including a film flag, a scene change flag, and a fade flag.

[0010] Moreover, this invention indicates the method of controlling the video encoder which encodes the input video data corresponding to two or more frames. The method of this invention contains the step to which only the predetermined number N of a frame ($N > 1$) delays an input video data, and generates a delay video data, the step which processes an input video data in order to generate the statistic of the 1st video frame, and the step which controls coding of the delay video data which is equivalent to the 1st frame with a video encoder using the aforementioned statistic.

[0011]

[Embodiments of the Invention] Before explaining this invention, a general matter is explained first. Since an image (picture) is constituted, the block of a pixel (pel) has been located in a line with the matrix. And each pixel relates to three elements, namely, brightness Y and the red difference Cr Blue difference Cb it is. The video data arranged in the field is national television system KAMITI "National Television System Committee" (NTSC). Standard video, it is about 59.94Hz (about 60Hz), and operates by 50Hz by the standard video for a phase ARUTA nation line "Phase Alternation Line" (PAL). A field pair constitutes a frame from a basis this NTSC standard and standard for PAL. FIELD1 calls the 1st field (odd number field) which should be displayed, and calls FIELD2 the 2nd field (even number field) which should be displayed. Each field is related to parity, i.e., even number, or odd number in this way.

[0012] Such the field is divided into a category again as the internal field (intra field) (I field), and the prediction field (P field) and the bidirectional field (B field), and these are similarly defined as I frames, P frames, and B frames.

[0013] In drawing 1, an encoder 10 has the encoder control module 11 which receives the command or other inputs from an input unit, the pre processor 12 which detects the film frame in a video input, and scene change, the rate controller 14, the encoder module 16, the prediction module 18, a formatter 20, the consciousness model module 22, the movement prediction module 24, and the decoder module 26. Conventionally, these component parts of an encoder 10 are well-known, for example, are indicated by U.S. Pat. No. 5,144,423, No. 5,231,484, No. 5,247,363, No. 5,293,229, No. 5,325,125, etc.

[0014] In the method indicated a video encoder control system and here, a pre processor 12 receives the input video data and command input which were processed by the encoder control module 11, and removes the redundant field from the video data corresponding to the film source. The rate controller 14 receives a data signal like the flag from a pre processor 12, and in order to perform coding, it controls operation of an encoder 10. The rate controller 14 controls communication with an external system and an encoder 10, in order to maintain the bit rate encoded in the bandwidth of operation again. The encoder module 16 receives the forecast from the video data and the prediction module 18 from a pre processor 12, and encodes the pretreated video data. A formatter 20 generates a coding output signal so that it can output through an output channel combining the block of the pixel of a video frame, and various data fields.

[0015] The consciousness model module 22 calculates a coding parameter for a coding process, and the movement prediction module 24 carries out block matching of the video data within the block of the present pixel, and the image data of precedence, moves, and generates a factor. The decoder module 26 generates the prediction error reconfigured from this coding process, and generates a decryption image.

[0016] In the example shown in drawing 2, a pre processor 12 has the look-up table module 28, it has this look-up table module 28 by the standard format of CCIR-601, receives an input video data, and performs operation of well-known arbitrary gamma corrections, pedestal adjustment, contrast emphasis, etc. Various tables are dedicated in the look-up table module 28 for the RUMA (luma) signal and the saturation (chroma) signal. The input video signal transmitted to the perpendicular cropping module 30 through the look-up table module 28 carries out the crop (generation) of the input video data.

[0017] For example, in order to process an input video data by the NTSC standard, the crop (generation) of this input video data is carried out to 480. On the other hand, in order to process an input video data by the standard [for PAL] one, the crop of this input video data is carried out to 576. Here, 576 lines of activity are used altogether. The perpendicular cropping module 30 is a high-definition TV standard "high definition television standards" (HDTV). The crop of the input video data is carried out according to the requirements standard [such as EGA, VGA, and super one VGA, / specific] for video. These standards are well-known. For example, the MPEG standard is indicated by "Coded Representation of Picture and Audio Information", ISO-IEC/JTC1/SC29/EG11, CCITT SG XV, Working Party XV/1, Document AVC-445b, and February 1993." of the MPEG test module 4.

[0018] the video data by which the crop was perpendicularly carried out to the video data from the NTSC source — the 3:2 pulldown processor 32 — receiving — PURUDAUN of 3:2 — detecting — an input data field like the redundant field — re — it processes in order to sequence-ize and to remove, and the video data of 24 per second is generated. This 3:2 pulldown processor 32 is indicated by U.S. Pat. No. 5,317,398 about this which performs removal of the formation of pulldown ** re-sequence of 3:2, and the field. In an example, this 3:2 pulldown processor 32 generates a delay video data. And drawing 3 has and explains other functions of this 3:2 pulldown processor 32.

[0019] In drawing 2, the delay video data from the 3:2 pulldown processor 32 receives with a perpendicular filter / sub sampler 34, and performs saturation subsampling (chroma subsampling). In this specification, subsampling shall include the sampling of a coefficient 1.

[0020] The encoder module 16 of an encoder 10 encodes the video data of 4:2:0 by the well-known method. Refer to U.S. Pat. No. 5,253,056 and No. 5,270,813 about this. This saturation channel resolution is the half of RUMA

resolution in both a horizontal direction and a perpendicular direction. To the input video data of a format of 4:2:2, a perpendicular filter / sub sampler 34 processes the saturation channel used by the encoder 10.

[0021] When this delay video data corresponds to a film or a non film, the 3:2 pulldown processor 32 shows a film state or a non film state with a film flag. A perpendicular filter / sub sampler 34 answers this film flag, the video data as a progressive film has delay film data with the frame basis using four predetermined tap filters, and it processes delay film data so that saturation filtering may be carried out. When it is the video data (interlaced) which jumped over this video data (therefore, it is not progressive film data), it has this video data by vertical saturation filtering independently to each field, and it is had and encoded at a perfect time rate. In this example, filtering of the odd number field is had and carried out with predetermined 7 tap filter, filtering of the even number field is had and carried out with predetermined 4 tap filter, and this predetermined tap filter is symmetrical.

[0022] This video data from a perpendicular filter / sub sampler 34 by which saturation filtering was carried out generates the video data which was processed by the level cropping module 40 the adapted type pre-filter 36, the level filter / sub sampler 38, and after that, and was pretreated, and this is outputted to the encoder module 16. Drawing 5 has and explains operation of the adapted type pre-filter 36, the level filter / sub sampler 38, and the level cropping module 40.

[0023] Next, in drawing 3, the 3:2 pulldown processor 32 detects the redundant field in the input video data by which the crop was carried out perpendicularly. It is used in case the film recorded by 24 per pulldown ** one second of 3:2 is displayed on the NTSC television system which operates by about 60Hz, and this is attained by displaying the mutual frame of a film in either for 1/20 second or 1/30 seconds. And this NTSC television camera records either the three fields or the two fields by the well-known method to each film frame, respectively.

[0024] The redundant field is generated in this pulldown process of 3:2. This 3:2 pulldown processor 32 receives an input video data, and when only the field N of a predetermined number (N > 1) delays an input video data, it generates a delay video data. Only the 1 field delays this input video data, respectively, and the 1st field delay 42 and the 2nd field delay 44 generate the redundant video data delayed the 2 fields. Then, this video data delayed the 2 fields is delayed by 8 field delay 46, a delay video data is generated, and this serves as an input video data delayed the 10 fields, i.e., five frames, in all.

[0025] The statistic generation machine 48 receives input data and the data by which only the 2 fields were delayed from the 2nd field delay 44, and generates a statistic from there.

[0026] the encoder module 16 processes a statistic with this inter-frame statistic by the delay video data (this example — five pieces — frame delay was carried out), and it enables it to adjust coding of consecutiveness of the delayed video data. Compression like the throat within a coding process can also be strengthened by having with a flag and a control signal like other data signals using this statistic, dividing a close video data into a category, and offering the look ahead function of an encoder 10. An encoder 10 is faced using such a statistic, a flag, etc., can remove the redundancy within an input video signal, before it encodes the comparatively difficult portion of a video image like change of the scene within a video image, it can predict it, and it can take prohibition action. According to a film, the flag from the pre processor 12 which expresses scene change, or other directions, an encoder 10 can schedule the following inner frame again from directions of the rate controller 14, it can have consistency, therefore coding image quality can be improved.

[0027] Setting in this example, an input video data is Y, Cb, and Cr. Corresponding to the field of the video pixel which it has, the statistic generation machine 48 calculates the following statistic from there to each input video field.

[0028] 1) The sum of the absolute value of the difference between the pixels of the input field which the same parity (even number, odd number) follows, [0029] 2) Maximum of the difference between the pixels of the image value to which low pass filter processing of the input field which the same parity follows (and the sub sample was formed when possible) was carried out. In this example, it is used as a sample of the image to which the block average of the value of a block of the pixel of input field was calculated, and the sub sample of the input field was carried out and by which low pass filter processing was carried out.

[0030] 3) The average of the pixel in input field.

[0031] This statistic generation machine 48 has a digital-signal-processing (DSP) circuit or a digital-signal-processing means like a chip, and this can use DSP1610 chip marketed from AT & T Corp. Furthermore, well-known digital-signal-processing software or other equivalent digital-signal-processing means are used with the statistic generation machine 48, and the above-mentioned statistic is determined. And this is used by the pulldown detector 50, the scene change detector 52, and the fade detector 53.

[0032] In drawing 4, the statistic generation machine 48 receives input data and 2 field delay-ized input data. This statistic generation machine 48 has the absolute value computer 54 of a difference, and generates the absolute value of the difference of input data and 2 field delay-ized input data, and the sum computer 56 of the field generates the sum of an absolute value covering the field using this absolute value. This sum is outputted to the pulldown detector 50 and the scene change detector 52 after that.

[0033] This statistic generation machine 48 has a low pass filter / sub samplers 58 and 60, and receives input data and 2 field delay-ized input data, respectively. The image value which was generated from the sum, which was formed into the sub sample and by which low pass filter processing was carried out is outputted to the absolute value computer 62 of a difference, and generates the absolute value of the difference between the image values by which low pass filter processing was carried out. The maximum detection machine 64 of the field covers the field, and determines the maximum absolute value, and this maximum absolute value is outputted to the pulldown detector 50.

[0034] A low pass filter / sub sampler 60 has the average computer 66 of this field covers the field, and determines the average of input data, and this average is outputted to the fade detector 53.

[0035] The following systems and method of the video encoder control system of this invention, and the following examples tend to detect film data scene change video phasing etc. using the statistic generation ratio from the input video data, and it is going to fit them to coding resolution.

[0036] In the 1st example of film detection, the pulldown detector 50 of the 3:2 pulldown processor 32 detects the film data which exist in an input video data using the sum of the absolute value of the value between the pixels between the fields of the alternation of an input video data. By the video data of the quiescence field or the redundant field (still or redundant field), the statistic generation machine 48 generates the low sum of the difference of an absolute value, and, on the other hand, the statistic generation machine 48 generates the sum (0 is approached) of the absolute value of a difference by the video data of a quiescence image (stationary image). Furthermore, when this video data is generated by conversion of 3:2 from a conversion film image, every 5 fields, the sum from the 5th field is small and, on the other hand, the remaining sums from the 4 fields of the 5 fields are large.

[0037] The pulldown detector 50 is detected by comparing the relative value of the sum generated from there in existence of the film in a video data, and generates the film flag with which a difference is taken out the 1st predetermined value and the pulldown detector 50 is transmitted to the rate controller 14 by this.

[0038] The pulldown detector 50 detects the redundant video field using the sum and the maximum absolute value of an absolute value of a difference of the image which was calculated by the statistic generation machine and by which low pass filter processing was carried out. Based on detection of the position of the redundant field, the group of the ten fields is distinguished as a film or a non film.

[0039] The sum of the absolute value of the difference between the input fields of a couple will express the relative mismatch between the fields of a couple with the MACROSS kale. When the sum of the absolute value of this difference is a small value, it is only the case where the redundant field exists in the field of a couple.

[0040] The maximum absolute value of the difference of the image by which low pass filter processing was carried out expresses the relative mismatch in the local field of the input field of this couple. When the maximum absolute value of this difference is small, it is only the case where the redundant field exists in the field of a couple.

[0041] When this input video data is generated from the film image changed by PURUDAUN of 3:2, the 5th field and the 10th field are the redundant fields among the groups of the ten fields. Furthermore, the grade of the mismatch between the 6th field and the 8th field bears a strong resemblance to the grade of the mismatch between the 7th field and the 9th field. The grade of the mismatch between the 4th field and the 6th field resembles the grade of the mismatch between the 5th field and the 7th field. The grade of the mismatch between the 1st field and the 3rd field resembles the grade of the mismatch between the 2nd field and the 4th field.

[0042] In this example, the pulldown detector 50 had six queues of FIFO (FIFO) of internal statistics of the length of eight units, and has memorized ten statistics which are the latest input field most. The queue is as follows.

[0043] a) {DY [0], ..., DY [7]} The sum of the absolute value of the field difference of a RUMA signal (luma signal) is memorized.

[0044] b) {DCr [0], ..., DCr [7]} memorizes the sum of the absolute value of the field difference of Cr saturation signal (chroma signal).

[0045] c) {DCb [0], ..., DCb [7]} is Cb. The sum of the absolute value of the field difference of a signal is memorized.

[0046] d) {dY [0], ..., dY [7]} memorizes the maximum absolute value of the field difference of the RUMA signal by which low pass filter processing was carried out.

[0047] e) {dCr [0], ..., dCr [7]} memorizes the maximum absolute value of the field difference of Cr saturation signal by which low pass filter processing was carried out.

[0048] f) {dCb [0], ..., dCb [7]} memorizes the maximum absolute value of the field difference of Cb saturation signal by which low pass filter processing was carried out.

[0049] In the above-mentioned explanation, the absolute value of the field difference over the predetermined field is calculated by the whole field by covering it.

[0050] The pulldown detector 50 holds a state variable psi again, and a state variable psi expresses the film mode of the field most outputted from 8 field delay 46 recently. When the field processed most recently is a non film, this state variable psi is defined as 0, and a state variable psi takes one value of the range of 1 to 10, and it expresses the sequence of the output field in 3:2 pulldown patterns of the 10 field. The state variable psi is first set to 0.

[0051] The following statistic is used for the pulldown detector 50 to each input field.

[0052] 1) The sum of the absolute value of the field difference of the pixel value between DY', the input brightness data about the field, and 2 field delay-ized brightness data [0053] 2) Cover DCr' and the field and it is the sum

[0054] of the absolute value of the field difference of the pixel value between input Cr saturation data and 2 field delay-ized Cr saturation data. 3) Cover DCb' and the field and it is the sum [0055] of the absolute value of the field difference of the pixel value between input Cb saturation data and 2 field delay-ized Cb saturation data. 4) Maximum of the absolute value of the field difference between the pixel values of the image which covers dY' and the field and by which low pass filter processing of input brightness data and the 2 field delay-ized brightness data was carried out [0056] 5) Maximum of the absolute value of the field difference between the pixel values of the image which covers dCr' and the field and by which low pass filter processing of input Cr saturation data and the 2 field delay-ized Cr saturation data was carried out [0057] 6) Maximum of the absolute value of the field difference between the pixel values of the image which covers dCb' and the field and by which low pass filter processing of

input Cb saturation data and the 2 field delay Cb saturation data was carried out [0058] This pulldown detector 50 updates FIFO of a statistic according to the following formula.

$DY[n] = DY[n-1]$

$DCr[n] = DCr[n-1]$

$DCb[n] = DCb[n-1]$

$dY[n] = dY[n-1]$

$dCr[n] = dCr[n-1]$

$dCb[n] = dCb[n-1]$

$n = 1, 2, \dots, 7$ — receiving — and $DY[0] = DY'$, $DCr[0] = DCr'$, $DCb[0] = DCb'$, $dY[0] = dY'$, $dCr[0] = dCr'$, $dCb[0] = dCb'$ [0059] When the present input field is [the present state of a state variable psi] 0 or 10 in the even (the

maximum bottom) fields, only 1 makes a state variable psi increase without the pulldown detector's 50 performing film detection as follows, or the pulldown detector 50 performing film detection.

[0060] The pulldown detector 50 determines that the following ten fields outputted from 8 field delay 46 only when all the states of formula (1)–(4) were fulfilled are films.

[0061]

[Equation 1]

[0062] Here, TY , TCr , TCb , tY , tCr , tCb , RY , RCr , RCb , RY' , RCr' , and RCb' are thresholds which were determined by the knowledge beforehand acquired from the input data, namely, exist beforehand.

[0063]

[Equation 2]

[0064] For example, these thresholds can be set up by a user's input command or training data.

[0065] When above all of formula (1)–(4) are filled (i.e., when it is shown that this film test of a film state is affirmative), a state variable psi is set as 1 and the class division of the following ten fields outputted from 8 field delay 46 is carried out as a film. Other than this, then, a state variable psi is set to 0, the following two fields outputted from 8 field delay 46 are classified as a non film (namely, video), and the above-mentioned detection process is repeated to new input data.

[0066] This pulldown detector 50 generates the number-of-letters value in the 2-bit mode again, and this has 0–3 (base 10), and shows reconstruction of a frame as shown in Table 1 to the rate controller 14.

[0067]

Table 1 film mode number-of-letters value Account ** 0 The field 1/the field 2 1 The field 2/the field 1 2

Recurrence of the field 1/the field 2/the field 1 3 Recurrence of the field 2/the field 1/the field 2 [0068] In the 2nd example of detection of change of a scene, the statistic generation machine 48 sends information to the pulldown detector 50, and this pulldown detector 50 detects the scene in input video, i.e., generating of change of the instant of a cut. The amount of change between the fields is classified using the sum of the absolute value of the difference between the pixels in the present field and 2 field delay-ized field. In the usual scene, even the case of the scene of the small movement which cannot be predicted covers several frames, and this parameter changes slowly. therefore, threshold $TLOW$ predetermined [scene] in all field differences in being the following, it considers that it is continuous — having — small-amount $\Delta TLOW$ only — it changes between continuous values

[0069] However, in the cut between two or more scenes, this parameter becomes large. The two fields used for the reason calculating a field difference are because it generates on a different scene. In this case, this field difference is usually larger than predetermined threshold $THIGH$. Furthermore, since a field difference is calculated between the mutual fields, when the cut of a scene occurs, the two continuous fields have a big field difference more than threshold $THIGH$.

[0070] Therefore, at least two high field differences can detect detection of change of a scene easily by detecting whether it comes after a low field difference, or it comes in front. Look ahead property of the statistic generation machine 48 (look-ahead nature) Detection of the number of consecutive field differences is advanced, and the reliability of detection of a scene is made to improve. Furthermore, since it is encoded, change of a scene is detected between 1 inputted into the encoder module 16, or two or more frame time. Although this which can change a coding parameter by this before scene change of an encoder 10 makes the coding quality of the frame before change of a scene deteriorate, since change of quality which is visible is not caused, thereby for those who are looking at the decrypted video, the coding efficiency of an encoder 10 is improvable.

[0071] If these change is detected, the scene change detector 52 will output the change flag of a scene, and will transmit this to the rate controller 14. For example, the change flag of this scene expresses a picture type like setting in coding. For example, to change of a scene which attains convinced level with detection, the flag in this frame is generated, it is sent to the rate controller 14, and internal coding (intra code) of the whole frame is carried out. It is outputted as the mutual resynchronization method from this flag scene change detector 52 in a frame. The statistic generation machine 48 has the counter (not shown) which generates a periodic count, and a pre processor 12 answers this periodic count, and generates the flag in a frame at a fixed interval. In this mutual resynchronization method, a pre processor 12 answers the change flag of a scene, generates the flag in a frame, and the rate controller 14 answers the flag in this frame, and it inserts an inner frame into the frame of the plurality of the place equivalent to change of a scene. In this method, according to the relative position of scene change, the scene change detector 52 is chosen so that the frequency of an inner frame may be changed, and it generates an inner frame, and makes scene change in agreement. And a pre processor 12 answers this scene change flag, corrects the count of a counter, and generates the flag in a frame at the time of correction of a count.

[0072] As an exception method, the scene change detector 52 outputs a 8-bit value to the rate controller 14, and expresses, the possibility, i.e., the strength, of scene change. This 8-bit value shows other scene change as partial scene change again.

[0073] In addition to detection of a film, the 3:2 pulldown processor 32 is possible also for inserting an inner frame at the predetermined intervals and inserting commercials and IP frame coding as another method, and successive per motion refreshment operation by this if it is attained.

[0074] The fade detector 53 puts for video phasing in a video data in the 3rd example of detection of video phasing. A twist generates the error detection of a film to the changes to the film added to the film from existence of the video fade and a film, or the fade rate of the changes to video from a film. This fade detector 53 detects changes of video from statistics, using an average pixel value as a thing showing a big change of the brightness of the scene before and behind the fade. This fade detector 53 answers that an average pixel value is below a predetermined value, and generates the fade flag showing the fade in video.

[0075] In the 4th example of adapted type resolution control, the video encoder control system of this invention fits the resolution of coding according to change of the scene measured by frame statistics. The statistic generation machine 48 of the 3:2 pulldown processor 32 generates the sum of the absolute value of the field difference of the pixel of an input video data as a scale of the corresponding complexity of an image. Change of the pixel between the fields is also calculated as an exception method. These sums are transmitted to the rate controller 14, and this rate controller 14 determines whether these sums exceed a predetermined value. If it becomes beyond this predetermined value (i.e., if the complexity of the picture which should be encoded is high), the rate controller 14 will generate the resolution selection flag on which the resolution of a coding image is dropped.

[0076] Although an input video data is processed by the adapted type pre-filter 36, in drawing 5, it is performed in this example using the 1st predetermined bank of the filter value of 256 pieces chosen by the rate controller 14, or a coefficient using a pre-filter selection signal at a rate of every frame 1 time. The adapted type pre-filter 36 receives a film flag from the 3:2 pulldown processor 32 again, and means whether this input video data is equivalent to either a progressive film or the video by which the interlace was carried out. This adapted type pre-filter 36 chooses a filter value from the 2nd predetermined bank of the group of a filter value using a film flag.

[0077] This film mode flag and a pre-filter selection signal are processed by 8x8 filters 68, and this 8x8 filter 68 consists of a group of eight level scan line taps and eight perpendicular taps, in order to perform filtering of 8x8 with the frame base. The tap of this 8x8 filter 68 is programmable. For example, if the tap of the matrix of an every other is set as 0, the filter of 8x8 turns into 4x4 filters, and in order to perform field filter processing, it can be adapted [this] for the pixel of a frame. The thing relevant to calculation of the output value to which this 8x8 filter 68 corresponds is only the value from one of the two fields per frame.

[0078] When this film flag is OFF, 8x8 filters 68 operate as a filter of 8x4, and perform filtering with the field base. And when a film flag is ON, 8x8 filters 68 perform filtering with the frame base, and the conditions (namely, state where the matrix in every other one of a tap is set as 0) concerning a tap are removed. In this way, to video, pre-filter processing is the field base and is performed with the frame base to a film.

[0079] The group of two or more filters, i.e., a filter value, or the group of a filter factor is given to filtering of each type, for example, filtering of the field base and filtering of the frame base. Such two or more filters can be determined by setting the tap of 8x8 filters 68 by different method. 8x8 filters 68 serve as an adapted type based on the degree of difficulty for encoding a video data based on complication of a video image, and the state of an encoder 10 by choosing from one of either the group of the value of a field filter, or the groups of the value of a frame filter besides suiting filtering of the field base or the frame base according to whether an input video data is film material or it is video material. The complexity of this scene is determined by the sum of the absolute value of the field difference of the continuation field of the same parity.

[0080] As an exception method, 8x8 filters 68 choose at least one of two or more of the filter values, in order to carry out filtering of the delay video data using the coding bit rate set up by the sum and the rate controller 14 of an absolute value of a field difference.

[0081] Since the thing with the predetermined group of this filter value is negative, the state of the overflow in filtering can avoid the output of 8x8 filters 68 by using a clamp 70, and clamping or latching it to 8-bit within the limits from 0 to 256.

[0082] Each Y and Cb Cr A value has with the filter value of a group which is different by the above-mentioned method, and filtering is independently carried out.

[0083] In this example, the adapted type pre-filter 36 outputs the data with which 720 pixels processed by the level filter / sub sampler 38 of 2:1 were clamped. This level filter / sub sampler 38 of 2:1 answer a resolution selection flag from the rate controller 14, and determines the rate of subsampling. The 704 mode and the 352 modes are supported in the example using the MPEG coding standard. When the 704 modes are chosen, a disable is carried out, the video data [finishing / processing] which has 704 pixels per line is inputted into the horizontal sampler 72 of 2:1, and the horizontal sampler 72 of 2:1 passes through that, without changing. If the 352 modes are chosen, they are Y, Cb, and Cr. By the coefficient of 2, a value has, sub sample processing is carried out, the 1st pixel of each line remains, and the resolution which is 352 pixels per line is offered.

[0084] In another example, it has a level filter / sub sampler 38 by the method same with performing frame filter processing for field filter processing, and it performs adapted type filtering using a film flag. In this way, this perpendicular filter / sub sampler 34 perform the down sampling of saturation, changes the saturation of 4:2:2 into the saturation of 4:2:0, and is decided by whether this is a film material as which the video data which should be

processed is determined by the 3:2 pulldown processor 32 with a film flag, or it is video material.

[0085]

[Effect of the Invention] Finer control of resolution is attained by using the resolution control method of this invention combining the above-mentioned adapted type pre-filter 36. In order to control the look ahead function and resolution of an encoder 10 from delay of an input video data furthermore, in case statistics are used, an encoder 10 The number of bits which resolution was raised using the resolution selection flag, and was assigned to at least one frame is decreased. In order to suppress generation of the artifact (imitation) of a coding process, use a resolution selection flag and resolution is made to increase before change of a scene, and it can respond to change of a scene by decreasing the number of bits assigned to at least one frame.

[0086] As shown in drawing 6, in case the method of this invention controls an encoder 10, it starts control of an encoder 10 at Step 74, it receives an input video data at Step 76, pretreats a video data at Step 78, and controls an encoder 10 by Step 88 using a delay video data and statistics. Only N frames contain in the step of pretreatment the step which delays a video data. A delay video data is generated at Step 80 here including N ($N > 1$). In order to generate a statistic from an input video data at Step 82 and to generate the flag which contains the change flag of a scene, a film flag, a video fade flag, etc. at Step 84, this statistic is processed, and it opts for a setup of resolution from this statistic at Step 86.

[Translation done.]

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

TECHNICAL FIELD

[The technical field to which invention belongs] About a video encoder, especially this invention detects the change in a video signal, and relates to the method and system which control the video encoder which performs coding.

[Translation done.]

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

PRIOR ART

[Description of the Prior Art] In the picture compression system, the video signal is digitized to the minimum number of bits, maintaining the greatest picture quality. The MPEG (Motion Picture Experts Group) standard has specified technology effective in picture compression. With the technology of a certain, the redundancy of data is maintained at the time of this compression, offering picture compression within a video encoder.

[0003] This well-known technology that has the film material of the input video image picture of 24 per for 1 second in a well-known process, and is changed into the video signal of the 60 fields per for 1 second is the pulldown method of 3:2, and this has each video frame in the three fields of video, and the two fields, and is recording it by turns. However, this pulldown method of 3:2 becomes redundant on the occasion of the conversion. Detecting such redundancy and removing it means removing the one field for every five fields, without causing informational loss. However, the video image data which records by this and should be processed can be reduced 20%. That no less than 20% of video image data can be reduced has the advantage which cuts down data capacity greatly to video coding, without losing image information, since 90% of television broadcasting material is obtained from the film source.

[0004] However, in order to detect such redundancy in the image field and to remove it, an artifact (imitation) to the extent that the detection which PURUDAUN of 3:2 in the mixture of a film and video or a different combination of a film mistook is nonpermissible on the occasion of processing of the compression image of consecutiveness like for example, an image decryption will be reconfigured.

[0005] It comes and prediction and prediction of corresponding correction of a coding parameter of change of ON video information are conventionally well-known. The quality is made to deteriorate in a consecutive decryption by such prediction method in the case by the arrangement which the inside of a frame (intra frames), i.e., a refreshment frame, mistook, or mistaken arrangement of I frames. For example, the usual video frame can usually be predicted from the frame in front of it at least, and in order to make it easy that a receiver acquires a video image, the inner flame (intra frames) transmitted periodically is attached to a receiver. Rather than the usual video frame, many bits are needed for coding of the frame from such coding in a frame, and change of a scene. The frame from which, as for the reason, such a scene changes is because it becomes the frame which can hardly be predicted. When generating in immediately after [just before coding in a frame changes / of a scene], an average bit rate required for coding will increase to the level to which the quality of coding is lowered, consequently an artifact (imitation) offensive to the eye will generate it on the occasion of a decryption.

[Translation done.]

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

EFFECT OF THE INVENTION

[Effect of the Invention] Finer control of resolution is attained by using the resolution control method of this invention combining the above-mentioned adapted type pre-filter 36. In order to control the look ahead function and resolution of an encoder 10 from delay of an input video data furthermore, in case statistics are used, an encoder 10 The number of bits which resolution was raised using the resolution selection flag, and was assigned to at least one frame is decreased. In order to suppress generation of the artifact (imitation) of a coding process, use a resolution selection flag and resolution is made to increase before change of a scene, and it can respond to change of a scene by decreasing the number of bits assigned to at least one frame.

[0086] As shown in drawing 6, in case the method of this invention controls an encoder 10, it starts control of an encoder 10 at Step 74, it receives an input video data at Step 76, pretreats a video data at Step 78, and controls an encoder 10 by Step 88 using a delay video data and statistics. Only N frames contain in the step of pretreatment the step which delays a video data. A delay video data is generated at Step 80 here including N ($N > 1$). In order to generate a statistic from an input video data at Step 82 and to generate the flag which contains the change flag of a scene, a film flag, a video fade flag, etc. at Step 84, this statistic is processed, and it opts for a setup of resolution from this statistic at Step 86.

[Translation done.]

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

TECHNICAL PROBLEM

[Problem(s) to be Solved by the Invention] Therefore, especially the purpose of this invention is detecting the change in a video signal and offering the efficient coding method of a frame signal.

[Translation done.]

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

MEANS

[Means for Solving the Problem] The processor of the system which controls the video encoder of this invention The multi-field delay circuit from which only a predetermined number N ($N > 1$) delays an input video data in order to generate a delay video data, The statistic generation machine which processes the aforementioned input video data in order to generate the statistic of the 1st frame and a consecutive frame and to generate a control signal from this statistic, It consists of an encoder module and a rate controller which controls coding of the delay video data which answers the aforementioned control signal and is equivalent to the 1st frame with the aforementioned encoder module. The aforementioned statistic generation machine calculates the sum of the absolute value of the difference between field pixels, and calculates the low pass filter image value by which the sub sample was carried out. And it is included in it, and using the aforementioned sum, a pulldown detector detects the redundant field in the related field, and generates a redundancy flag as a control signal corresponding to the redundant field.

[0008] It is included in it, and using the aforementioned sum, a scene change detector detects change of the scene from the 1st frame and a consecutive frame, and generates the change flag of this scene as a control signal. A statistic generation machine determines the average pixel value of each frame, and in order to determine the video fade, the average pixel value of the 1st field and the consecutive field is used for a fade detector.

[0009] A resolution selection signal is generated using the aforementioned sum, a rate controller uses this, and a resolution selector determines the resolution of coding of a consecutive frame. The output from a processor is sent to a rate controller, in order to control coding of a delay video data including a film flag, a scene change flag, and a fade flag.

[0010] Moreover, this invention indicates the method of controlling the video encoder which encodes the input video data corresponding to two or more frames. The method of this invention contains the step to which only the predetermined number N of a frame ($N > 1$) delays an input video data, and generates a delay video data, the step which processes an input video data in order to generate the statistic of the 1st video frame, and the step which controls coding of the delay video data which is equivalent to the 1st frame with a video encoder using the aforementioned statistic.

[0011]

[Embodiments of the Invention] Before explaining this invention, a general matter is explained first. Since an image (picture) is constituted, the block of a pixel (pel) has been located in a line with the matrix. And each pixel relates to three elements, namely, brightness Y and the red difference Cr Blue difference Cb it is. The video data arranged in the field is national television system KAMITI "National Television System Committee" (NTSC). Standard video, it is about 59.94Hz (about 60Hz), and operates by 50Hz by the standard video for a phase ARUTA nation line "Phase Alternation Line" (PAL). A field pair constitutes a frame from a basis this NTSC standard and standard for PAL. FIELD1 calls the 1st field (odd number field) which should be displayed, and calls FIELD2 the 2nd field (even number field) which should be displayed. Each field is related to parity, i.e., even number, or odd number in this way.

[0012] Such the field is divided into a category again as the internal field (intra field) (I field), and the prediction field (P field) and the bidirectional field (B field), and these are similarly defined as I frames, P frames, and B frames.

[0013] In drawing 1, an encoder 10 has the encoder control module 11 which receives the command or other inputs from an input unit, the pre processor 12 which detects the film frame in a video input, and scene change, the rate controller 14, the encoder module 16, the prediction module 18, a formatter 20, the consciousness model module 22, the movement prediction module 24, and the decoder module 26. Conventionally, these component parts of an encoder 10 are well-known, for example, are indicated by U.S. Pat. No. 5,144,423, No. 5,231,484, No. 5,247,363, No. 5,293,229, No. 5,325,125, etc.

[0014] In the method indicated a video encoder control system and here, a pre processor 12 receives the input video data and command input which were processed by the encoder control module 11, and removes the redundant field from the video data corresponding to the film source. The rate controller 14 receives a data signal like the flag from a pre processor 12, and in order to perform coding, it controls operation of an encoder 10. The rate controller 14 controls communication with an external system and an encoder 10, in order to maintain the bit rate encoded in the bandwidth of operation again. The encoder module 16 receives the forecast from the video data and the prediction module 18 from a pre processor 12, and encodes the pretreated video data. A formatter 20 generates a coding output signal so that it can output through an output channel combining the block of the pixel of a video frame, and various data fields.

[0015] The consciousness model module 22 calculates a coding parameter for a coding process, and the movement prediction module 24 carries out block matching of the video data within the block of the present pixel, and the

image data of precedence, motion vectors, and generation as a factor. The decoder module 26 generates the prediction error reconfigured from this coding process, and generates a decoded image.

[0016] In the example shown in drawing 2, a preprocessor 12 has the look-up table module 28, it has this look-up table module 28 by the standard format of CCIR-601, receives an input video data, and performs operation of well-known arbitrary gamma corrections, pedestal adjustment, contrast emphasis, etc. Various tables are dedicated in the look-up table module 28 for the YUMA (luma) signal and the saturation (chroma) signal. The input video signal transmitted to the perpendicular cropping module 30 through the look-up table module 28 carries out the crop (generation) of the input video data.

[0017] For example, in order to process an input video data by the NTSC standard, the crop (generation) of this input video data is carried out to 480. On the other hand, in order to process an input video data by the standard [for PAL] one, the crop of this input video data is carried out to 576. Here, 576 lines of activity are used altogether. The perpendicular cropping module 30 is a high-definition TV standard "high definition television standards" (HDTV). The crop of the input video data is carried out according to the requirements standard [such as EGA, VGA, and super one VGA, / specific] for video. These standards are well-known. For example, the MPEG standard is indicated by "Coded Representation of Picture and Audio Information", ISO-IEC/JTC1/SC29/EG11, CCITT SG XV, Working Party XV/1, Document AVC-445b, and February 1993." of the MPEG test module 4.

[0018] The video data by which the crop was perpendicularly carried out to the video data from the NTSC source — the 3:2 pulldown processor 32 — receiving — PURUDAWN of 3:2 — detecting — an input data field like the redundant field — re—— it processes in order to sequence-size and to remove, and the video data of 24 per second is generated. This 3:2 pulldown processor 32 is indicated by U.S. Pat. No. 5,317,398 about this which performs removal of the formation of pulldown ** re-sequence of 3:2, and the field. In an example, this 3:2 pulldown processor 32 generates a delay video data. And drawing 3 has and explains other functions of this 3:2 pulldown processor 32.

[0019] In drawing 2, the delay video data from the 3:2 pulldown processor 32 receives with a perpendicular filter / sub sampler 34, and performs saturation subsampling (chroma subsampling). In this specification, subsampling shall include the sampling of a coefficient 1.

[0020] The encoder module 16 of an encoder 10 encodes the video data of 4:2:0 by the well-known method. Refer to U.S. Pat. No. 5,253,056 and No. 5,270,813 about this. This saturation channel resolution is the half of YUMA resolution in both a horizontal direction and a perpendicular direction. To the input video data of a format of 4:2:2, a perpendicular filter / sub sampler 34 processes the saturation channel used by the encoder 10.

[0021] When this delay video data corresponds to a film or a non film, the 3:2 pulldown processor 32 shows a film state or a non film state with a film flag. A perpendicular filter / sub sampler 34 answers this film flag, the video data as a progressive film has delay film data with the frame base using four predetermined tap filters, and it processes delay film data so that saturation filtering may be carried out. When it is the video data (interlaced) which jumped over this video data (therefore, it is not progressive film data), it has this video data by vertical saturation filtering independently to each field, and it is had and encoded at a perfect time rate. In this example, filtering of the odd number field is had and carried out with predetermined 7 tap filter, filtering of the even number field is had and carried out with predetermined 4 tap filter, and this predetermined tap filter is symmetrical.

[0022] This video data from a perpendicular filter / sub sampler 34 by which saturation filtering was carried out generates the video data which was processed by the level cropping module 40 the adapted type pre-filter 36, the level filter / sub sampler 38, and after that, and was pretreated, and this is outputted to the encoder module 16. Drawing 5 has and explains operation of the adapted type pre-filter 36, the level filter / sub sampler 38, and the level cropping module 40.

[0023] Next, in drawing 3, the 3:2 pulldown processor 32 detects the redundant field in the input video data by which the crop was carried out perpendicularly. It is used in case the film recorded by 24 per pulldown ** one second of 3:2 is displayed on the NTSC television system which operates by about 60Hz, and this is attained by displaying the mutual frame of a film in either for 1/20 second or 1/30 seconds. And this NTSC television camera records either the three fields or the two fields by the well-known method to each film frame, respectively.

[0024] The redundant field is generated in this pulldown process of 3:2. This 3:2 pulldown processor 32 receives an input video data, and when only the field N of a predetermined number ($N > 1$) delays an input video data, it generates a delay video data. Only the 1 field delays this input video data, respectively, and the 1st field delay 42 and the 2nd field delay 44 generate the redundant video data delayed the 2 fields. Then, this video data delayed the 2 fields is delayed by 8 field delay 46, a delay video data is generated, and this serves as an input video data delayed the 10 fields, i.e., five frames, in all.

[0025] The statistic generation machine 48 receives input data and the data by which only the 2 fields were delayed from the 2nd field delay 44, and generates a statistic from there.

[0026] the encoder module 16 processes a statistic with this inter-frame statistic by the delay video data (this example — five pieces — frame delay was carried out), and it enables it to adjust coding of consecutiveness of the delayed video data. Compression like the throat within a coding process can also be strengthened by having with a flag and a control signal like other data signals using this statistic, dividing a close video data into a category, and off ring the look ahead function of an encoder 10. An encoder 10 is faced using such a statistic, a flag, etc., can remove the redundancy within an input video signal, but it encodes the comparatively difficult portion of a video image like change of the scene within a video image, it can predict it, and it can take prohibition action. According to a film, the flag from the pre processor 12 which expresses scene change, or other directions, an encoder 10 can schedule-size the following inner frame again from directions of the rate controller 14, it can have consistency,

ther for coding image quality can be improved.

[0027] Setting in this example, an input video data is Y, Cb, and Cr. Corresponding to the field of the video pixel which it has, the statistic generation machine 48 calculates the following statistic from the input video field.

[0028] 1) The sum of the absolute value of the difference between the pixels of the input field which the same parity (even number, odd number) follows, [0029] 2) Maximum of the difference between the pixels of the image value to which low pass filter processing of the input field which the same parity follows (and the sub sample was formed when possible) was carried out. In this example, it is used as a sample of the image to which the block average of the value of a block of the pixel of input field was calculated, and the sub sample of the input field was carried out and by which low pass filter processing was carried out.

[0030] 3) The average of the pixel in input field.

[0031] This statistic generation machine 48 has a digital-signal-processing (DSP) circuit or a digital-signal-processing means like a chip, and this can use DSP1610 chip marketed from AT & T Corp. Furthermore, well-known digital-signal-processing software or other equivalent digital-signal-processing means are used with the statistic generation machine 48, and the above-mentioned statistic is determined. And this is used by the pulldown detector 50, the scene change detector 52, and the fade detector 53.

[0032] In drawing 4, the statistic generation machine 48 receives input data and 2 field delay-ized input data. This statistic generation machine 48 has the absolute value computer 54 of a difference, and generates the absolute value of the difference of input data and 2 field delay-ized input data, and the sum computer 56 of the field generates the sum of an absolute value covering the field using this absolute value. This sum is outputted to the pulldown detector 50 and the scene change detector 52 after that.

[0033] This statistic generation machine 48 has a low pass filter / sub samplers 58 and 60, and receives input data and 2 field delay-ized input data, respectively. The image value which was generated from these, which was formed into the sub sample and by which low pass filter processing was carried out is outputted to the absolute value computer 62 of a difference, and generates the absolute value of the difference between the image values by which low pass filter processing was carried out. The maximum detection machine 64 of the field covers the field, and determines the maximum absolute value, and this maximum absolute value is outputted to the pulldown detector 50.

[0034] A low pass filter / sub sampler 60 has the average computer 66 of the field again, the average computer 66 of this field covers the field, and determines the average of input data, and this average is outputted to the fade detector 53.

[0035] The following systems and method of the video encoder control system of this invention, and the following examples tend to detect film data scene change video phasing etc. using the statistic generated from the input video data, and it is going to fit them to coding resolution.

[0036] In the 1st example of film detection, the pulldown detector 50 of the 3:2 pulldown processor 32 detects the film data which exist in an input video data using the sum of the absolute value of the value between the pixels between the fields of the alternation of an input video data. By the video data of the quiescence field or the redundant field (still or redundant field), the statistic generation machine 48 generates the low sum of the difference of an absolute value, and, on the other hand, the statistic generation machine 48 generates the sum (0 is approached) of the absolute value of a difference by the video data of a quiescence image (stationary image). Furthermore, when this video data is generated by conversion of 3:2 from a conversion film image, every 5 fields, the sum from the 5th field is small and, on the other hand, the remaining sums from the 4 fields of the 5 fields are large.

[0037] The pulldown detector 50 is detected by comparing the relative value of the sum generated from there in existence of the film in a video data, and generates the film flag with which a difference is taken out the 1st predetermined value and the pulldown detector 50 is transmitted to the rate controller 14 by this.

[0038] The pulldown detector 50 detects the redundant video field using the sum and the maximum absolute value of an absolute value of a difference of the image which was calculated by the statistic generation machine and by which low pass filter processing was carried out. Based on detection of the position of the redundant field, the group of the ten fields is distinguished as a film or a non film.

[0039] The sum of the absolute value of the difference between the input fields of a couple will express the relative mismatch between the fields of a couple with the MACROSS kale. When the sum of the absolute value of this difference is a small value, it is only the case where the redundant field exists in the field of a couple.

[0040] The maximum absolute value of the difference of the image by which low pass filter processing was carried out expresses the relative mismatch in the local field of the input field of this couple. When the maximum absolute value of this difference is small, it is only the case where the redundant field exists in the field of a couple.

[0041] When this input video data is generated from the film image changed by PURUDAUN of 3:2, the 5th field and the 10th field are the redundant fields among the groups of the ten fields. Furthermore, the grade of the mismatch between the 6th field and the 8th field bears a strong resemblance to the grade of the mismatch between the 7th field and the 9th field. The grade of the mismatch between the 4th field and the 6th field resembles the grade of the mismatch between the 5th field and the 7th field. The grade of the mismatch between the 1st field and the 3rd field resembles the grade of the mismatch between the 2nd field and the 4th field.

[0042] In this example, the pulldown detector 50 had six queues of FIFO (FIFO) of internal statistics of the length of eight units, and has memorized ten statistics which are the latest input field most. The queue is as follows.

[0043] a) [DY [0], ..., DY [7]] The sum of the absolute value of the field difference of a RUMA signal (luma signal) is memorized.

[0044] b) [DCr [0], ..., DCr [7]] memorizes the sum of the absolute value of the field difference of Cr saturation signal (chroma signal).

[0045] c) [DCb [0], ..., DCb [7]] is Cb. The sum of the absolute value of the field difference of a signal is memorized.

[0046] d) [dY [0], ..., dY [7]] memorizes the maximum absolute value of the field difference of the RUMA signal by which low pass filter processing was carried out.

[0047] e) [dCr [0], ..., dCr [7]] memorizes the maximum absolute value of the field difference of Cr saturation signal by which low pass filter processing was carried out.

[0048] f) [dCb [0], ..., dCb [7]] memorizes the maximum absolute value of the field difference of Cb saturation signal by which low pass filter processing was carried out.

[0049] In the above-mentioned explanation, the absolute value of the field difference over the predetermined field is calculated by the whole field by covering it.

[0050] The pulldown detector 50 holds a state variable psi again, and a state variable psi expresses the film mode of the field most outputted from 8 field delay 46 recently. When the field processed most recently is a non film, this state variable psi is defined as 0, and a state variable psi takes one value of the range of 1 to 10, and it expresses the sequence of the output field in 3:2 pulldown patterns of the 10 field. The state variable psi is first set to 0.

[0051] The following statistic is used for the pulldown detector 50 to each input field.

[0052] 1) The sum of the absolute value of the field difference of the pixel value between DY', the input brightness data about the field, and 2 field delay-ized brightness data [0053] 2) Cover DCr' and the field and it is the sum [0054] of the absolute value of the field difference of the pixel value between input Cr saturation data and 2 field delay-ized Cr saturation data. 3) Cover DCb' and the field and it is the sum [0055] of the absolute value of the field difference of the pixel value between input Cb saturation data and 2 field delay-ized Cb saturation data. 4) Maximum of the absolute value of the field difference between the pixel values of the image which covers dY' and the field and by which low pass filter processing of input brightness data and the 2 field delay-ized brightness data was carried out [0056] 5) Maximum of the absolute value of the field difference between the pixel values of the image which covers dCr' and the field and by which low pass filter processing of input Cr saturation data and the 2 field delay-ized Cr saturation data was carried out [0057] 6) Maximum of the absolute value of the field difference between the pixel values of the image which covers dCb' and the field and by which low pass filter processing of input Cb saturation data and the 2 field delay-ized Cb saturation data was carried out [0058] This pulldown detector 50 updates FIFO of a statistic according to the following formula.

$DY[n]=DY[n-1]$

$DCr[n]=DCr[n-1]$

$DCb[n]=DCb[n-1]$

$dY[n]=dY[n-1]$

$dCr[n]=dCr[n-1]$

$dCb[n]=dCb[n-1]$

$n = 1, 2, \dots, 7$ — receiving — and $DY[0] = DY' \cdot DCr[0] = DCr' \cdot DCb[0] = DCb' \cdot dY[0] = dY' \cdot dCr[0] = dCr' \cdot dCb[0]$

$= dCb'$ [0059] When the present input field is [the present state of a state variable psi] 0 or 10 in the even (the maximum bottom) fields, only 1 makes a state variable psi increase without the pulldown detector's 50 performing film detection as follows, or the pulldown detector 50 performing film detection.

[0060] The pulldown detector 50 determines that the following ten fields outputted from 8 field delay 46 only when all the states of formula (1)–(4) were fulfilled are films.

[0061]

[Equation 1]

[0062] Here, TY, TCr, TCb, tY, tCr, tCb, RY, RCr, RCb, RY', RCr', and RCb' are thresholds which were determined by the knowledge beforehand acquired from the input data, namely, exist beforehand.

[0063]

[Equation 2]

[0064] For example, these thresholds can be set up by a user's input command or training data.

[0065] When above all of formula (1)–(4) are filled (i.e., when it is shown that this film test of a film state is affirmative), a state variable psi is set as 1 and the class division of the following ten fields outputted from 8 field delay 46 is carried out as a film. Other than this, then, a state variable psi is set to 0, the following two fields outputted from 8 field delay 46 are classified as a non film (namely, video), and the above-mentioned detection process is repeated to new input data.

[0066] This pulldown detector 50 generates the number-of-letters value in the 2-bit mode again, and this has 0–3 (base 10), and shows reconstruction of a frame as shown in Table 1 to the rate controller 14.

[0067]

Table 1 film mode number-of-letters value Account ** 0 The field 1/the field 2 1 The field 2/the field 1 2

Recurrence of the field 1/the field 2/the field 1 3 Recurrence of the field 2/the field 1/the field 2 [0068] In the 2nd example of detection of change of a scene, the statistic generation machine 48 sends information to the pulldown detector 50, and this pulldown detector 50 detects the scene in input video, i.e., generating of change of the instant of a cut. The amount of change between the fields is classified using the sum of the absolute value of the difference between the pixels in the present field and 2 field delay-ized field. In the usual scene, even the case of the scene of the small movement which cannot be predicted covers several frames, and this parameter changes slowly. therefore, threshold TLOW predetermined [scene] in all field differences in being the following, it considers that it is

continuous — having — small amount deltaTLOW only — it changes between continuous values

[0069] However, in the cut between two remote scenes, this parameter becomes large. The two fields used for the reason calculating a field difference are because it generates on a different scene. In this case, this field difference is usually larger than predetermined threshold THIGH. Furthermore, since a field difference is calculated between the mutual fields, when the cut of a scene occurs, the two continuous fields have a big field difference more than threshold THIGH.

[0070] Therefore, at least two high field differences can detect detection of change of a scene easily by detecting whether it comes after a low field difference, or it comes in front. Look ahead property of the statistic generation machine 48 (look-ahead nature) Detection of the number of consecutive field differences is advanced, and the reliability of detection of a scene is made to improve. Furthermore, since it is encoded, change of a scene is detected between 1 inputted into the encoder module 16, or two or more frame time. Although this which can change a coding parameter by this before scene change of an encoder 10 makes the coding quality of the frame before change of a scene deteriorate, since change of quality which is visible is not caused, thereby for those who are looking at the decrypted video, the coding efficiency of an encoder 10 is improvable.

[0071] If these change is detected, the scene change detector 52 will output the change flag of a scene, and will transmit this to the rate controller 14. For example, the change flag of this scene expresses a picture type like setting in coding. For example, to change of a scene which attains convinced level with detection, the flag in this frame is generated, it is sent to the rate controller 14, and internal coding (intra code) of the whole frame is carried out. It is outputted as the mutual resynchronization method from this flag scene change detector 52 in a frame. The statistic generation machine 48 has the counter (not shown) which generates a periodic count, and a pre processor 12 answers this periodic count, and generates the flag in a frame at a fixed interval. In this mutual resynchronization method, a pre processor 12 answers the change flag of a scene, generates the flag in a frame, and the rate controller 14 answers the flag in this frame, and it inserts an inner frame into the frame of the plurality of the place equivalent to change of a scene. In this method, according to the relative position of scene change, the scene change detector 52 is chosen so that the frequency of an inner frame may be changed, and it generates an inner frame, and makes scene change in agreement. And a pre processor 12 answers this scene change flag, corrects the count of a counter, and generates the flag in a frame at the time of correction of a count.

[0072] As an exception method, the scene change detector 52 outputs a 8-bit value to the rate controller 14, and expresses, the possibility, i.e., the strength, of scene change. This 8-bit value shows other scene change as partial scene change again.

[0073] In addition to detection of a film, the 3:2 pulldown processor 32 is possible also for inserting an inner frame at the predetermined intervals and inserting commercials and IP frame coding as another method, and successive promotion refreshment operation by this of it is attained.

[0074] The fade detector 53 opts for video phasing in a video data in the 3rd example of detection of video phasing. A twist generates the error detection of a film to the changes to the film added to the film from existence of the video fade and a film, or the fade rate of the changes to video from a film. This fade detector 53 detects changes of video from statistics, using an average pixel value as a thing showing a big change of the brightness of the scene before and behind the fade. This fade detector 53 answers that an average pixel value is below a predetermined value, and generates the fade flag showing the fade in video.

[0075] In the 4th example of adapted type resolution control, the video encoder control system of this invention fits the resolution of coding according to change of the scene measured by frame statistics. The statistic generation machine 48 of the 3:2 pulldown processor 32 generates the sum of the absolute value of the field difference of the pixel of an input video data as a scale of the corresponding complexity of an image. Change of the pixel between the fields is also calculated as an exception method. These sums are transmitted to the rate controller 14, and this rate controller 14 determines whether these sums exceed a predetermined value. If it becomes beyond this predetermined value (i.e., if the complexity of the picture which should be encoded is high), the rate controller 14 will generate the resolution selection flag on which the resolution of a coding image is dropped.

[0076] Although an input video data is processed by the adapted type pre-filter 36, in drawing 5, it is performed in this example using the 1st predetermined bank of the filter value of 256 pieces chosen by the rate controller 14, or a coefficient using a pre-filter selection signal at a rate of every frame 1 time. The adapted type pre-filter 36 receives a film flag from the 3:2 pulldown processor 32 again, and means whether this input video data is equivalent to either a progressive film or the video by which the interlace was carried out. This adapted type pre-filter 36 chooses a filter value from the 2nd predetermined bank of the group of a filter value using a film flag.

[0077] This film mode flag and a pre-filter selection signal are processed by 8x8 filters 68, and this 8x8 filter 68 consists of a group of eight level scan line taps and eight perpendicular taps, in order to perform filtering of 8x8 with the frame base. The tap of this 8x8 filter 68 is programmable. For example, if the tap of the matrix of an every other is set as 0, the filter of 8x8 turns into 4x4 filters, and in order to perform field filter processing, it can be adapted [this] for the pixel of a frame. The thing relevant to calculation of the output value to which this 8x8 filter 68 corresponds is only the value from one of the two fields per frame.

[0078] When this film flag is OFF, 8x8 filters 68 operate as a filter of 8x4, and perform filtering with the field base. And when a film flag is ON, 8x8 filters 68 perform filtering with the frame base, and the conditions (namely, state where the matrix in every other one of a tap is set as 0) concerning a tap are removed. In this way, to video, pre-filter processing is the field base and is performed with the frame base to a film.

[0079] The group of two or more filters, i.e., a filter value, the group of a filter factor is given to filtering of each

type, for example, filtering of the field basis and filtering of the frame basis. Such two or more filters can be determined by setting the tap of 8x8 filters 68 by different method. 8x8 filters 68 serve as an adapted type based on the degree of difficulty for encoding a video data based on complication of a video image, and the state of an encoder 10 by choosing from one of either the group of the value of a field filter, or the groups of the value of a frame filter besides suiting filtering of the field basis or the frame basis according to whether an input video data is film material or it is video material. The complexity of this scene is determined by the sum of the absolute value of the field difference of the continuation field of the same parity.

[0080] As an exception method, 8x8 filters 68 choose at least one of two or more of the filter values, in order to carry out filtering of the delay video data using the coding bit rate set up by the sum and the rate controller 14 of an absolute value of a field difference.

[0081] Since the thing with the predetermined group of this filter value is negative, the state of the overflow in filtering can avoid the output of 8x8 filters 68 by using a clamp 70, and clamping or latching it to 8-bit within the limits from 0 to 256.

[0082] Each Y and Cb Cr A value has with the filter value of a group which is different by the above-mentioned method, and filtering is independently carried out.

[0083] In this example, the adapted type pre-filter 36 outputs the data with which 720 pixels processed by the level filter / sub sampler 38 of 2:1 were clamped. This level filter / sub sampler 38 of 2:1 answers a resolution selection flag from the rate controller 14, and determines the rate of subsampling. The 704 mode and the 352 modes are supported in the example using the MPEG coding standard. When the 704 modes are chosen, a disable is carried out, the video data [finishing / processing] which has 704 pixels per line is inputted into the horizontal sampler 72 of 2:1, and the horizontal sampler 72 of 2:1 passes through that, without changing. If the 352 modes are chosen, they are Y, Cb, and Cr. By the coefficient of 2, a value has, sub sample processing is carried out, the 1st pixel of each line remains, and the resolution which is 352 pixels per line is offered.

[0084] In another example, it has a level filter / sub sampler 38 by the method same with performing frame filter processing or field filter processing, and it performs adapted type filtering using a film flag. In this way, this perpendicular filter / sub sampler 34 perform the down sampling of saturation, changes the saturation of 4:2:2 into the saturation of 4:2:0, and is decided by whether this is a film material as which the video data which should be processed is determined by the 3:2 pulldown processor 32 with a film flag, or it is video material.

[Translation done.]

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] The block diagram of the video encoder of this invention

[Drawing 2] The block diagram of a pre processor

[Drawing 3] The block diagram of the pulldown processor of 3:2 of drawing 2

[Drawing 4] The block diagram of the statistic generation machine of the pulldown processor of 3:2 of drawing 2

[Drawing 5] The block diagram showing the element of the pre processor of drawing 2

[Drawing 6] The flow chart view showing the method of the video encoder control system of this invention, and operation

[Description of Notations]

10 Encoder

11 Encoder Control Module

12 Pre Processor

14 Rate Controller

16 Encoder Module

18 Prediction Module

20 Formatter

22 Consciousness Model Module

24 Movement Prediction Module

26 Decoder Module

28 Look-up Table Module

30 Perpendicular Cropping Module

32 3:2 Pulldown Processor

34 Perpendicular Filter / Sub Sampler

36 Adapted Type Pre-filter

38 Level Filter / Sub Sampler

40 Level Cropping Module

42 1st Field Delay

44 2nd Field Delay

46 8 Field Delay

48 Statistic Generation Machine

50 Pulldown Detector

52 Scene Change Detector

53 Fade Detector

54 62 Absolute value computer of a difference

56 Sum Computer of Field

58 60 A low pass filter / sub sampler

64 Maximum Detection Machine of Field

66 Average Computer of Field

68 8X8 Filters

70 Clamp

72 2:1 Horizontal Sampler

74 Control Start of Video Encoder

76 Receive Input Video Data.

78 Pretreatment of Video Data

80 In Order to Generate Delay Video Data, Delay Only N Video Datas.

82 Generate Statistic from Input Video Data.

84 Generate Flag Which Processes this Statistic and Contains Scene Change Flag, Film Flag, and Video Fade Flag.

86 Opt for Step up of Resolution from this Statistic.

88 Control Encoder Using Delay Video Data and Statistic.

[Translation done.]

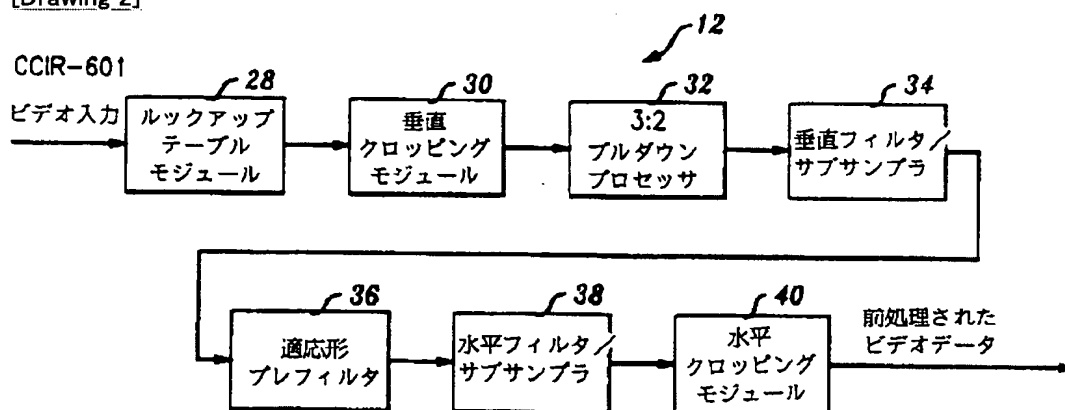
* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

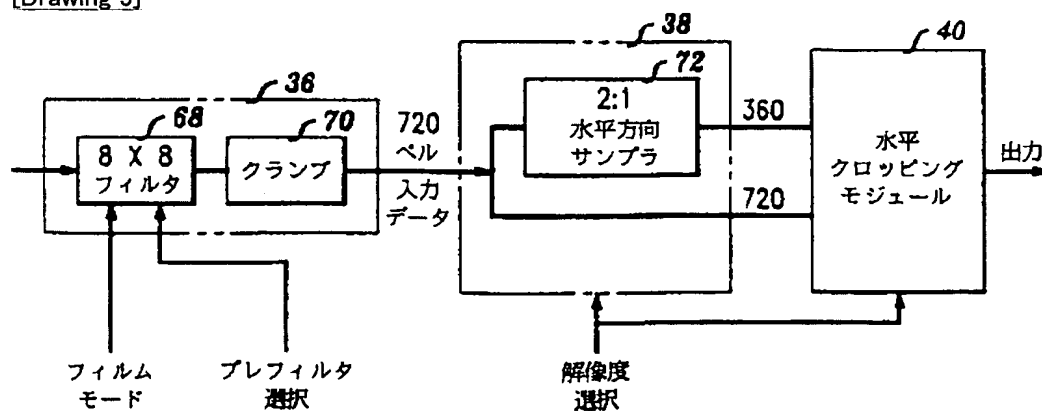
1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DRAWINGS

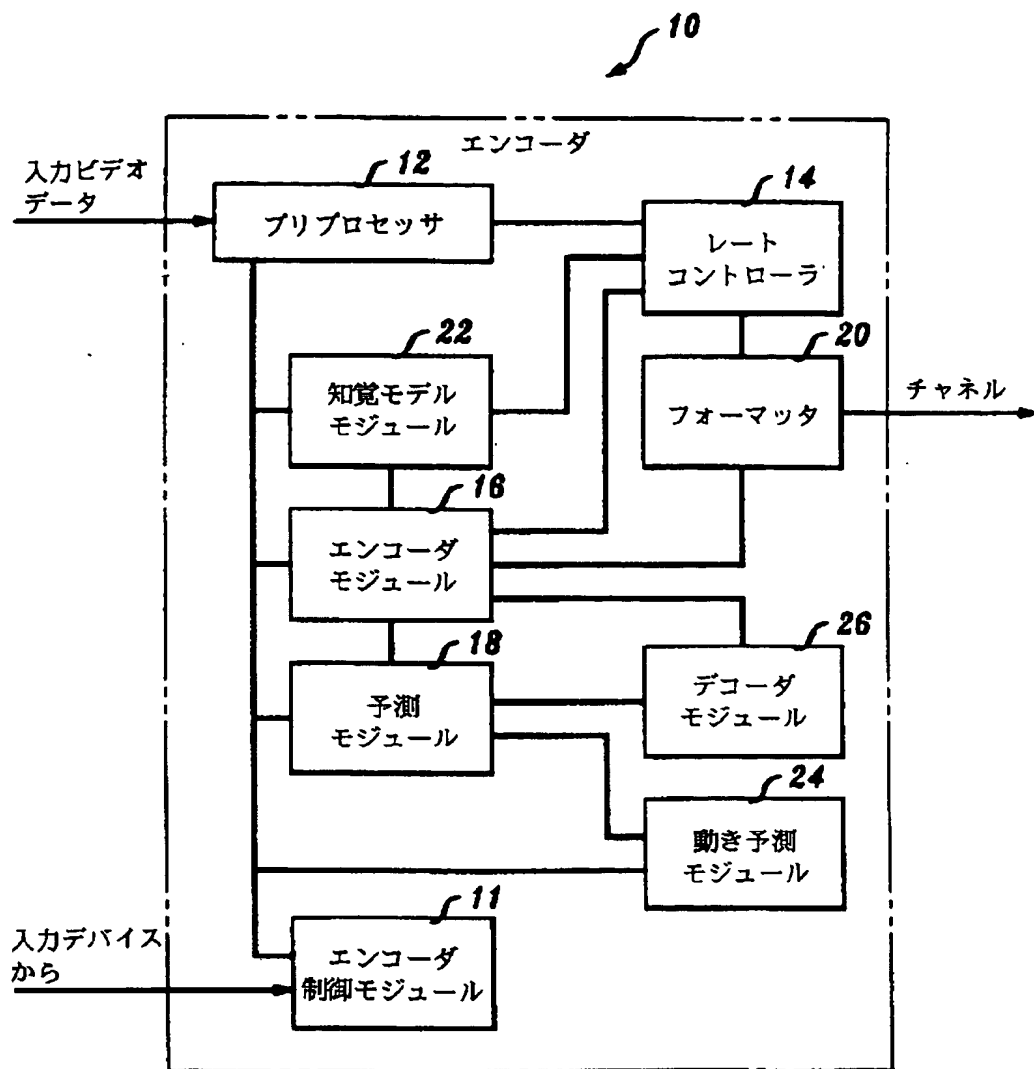
[Drawing 2]



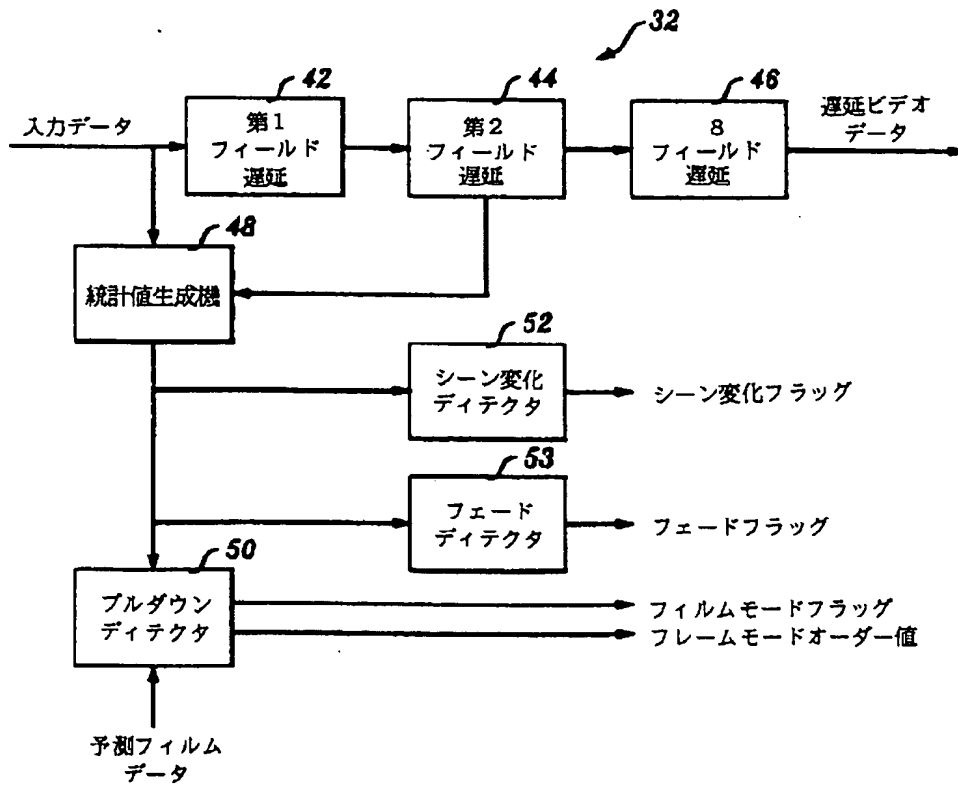
[Drawing 5]



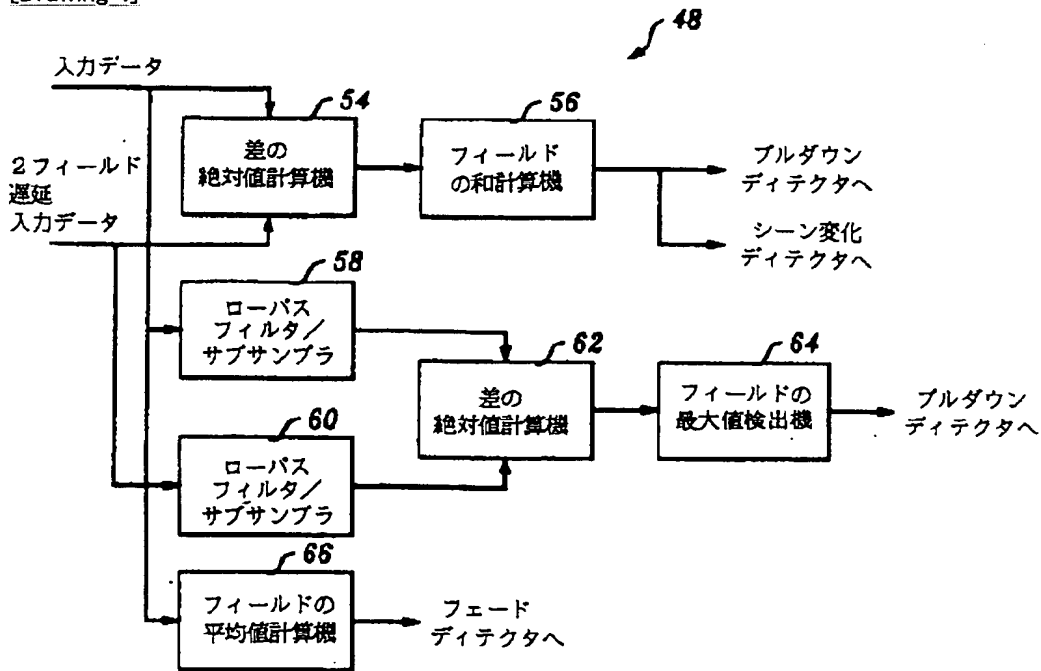
[Drawing 1]



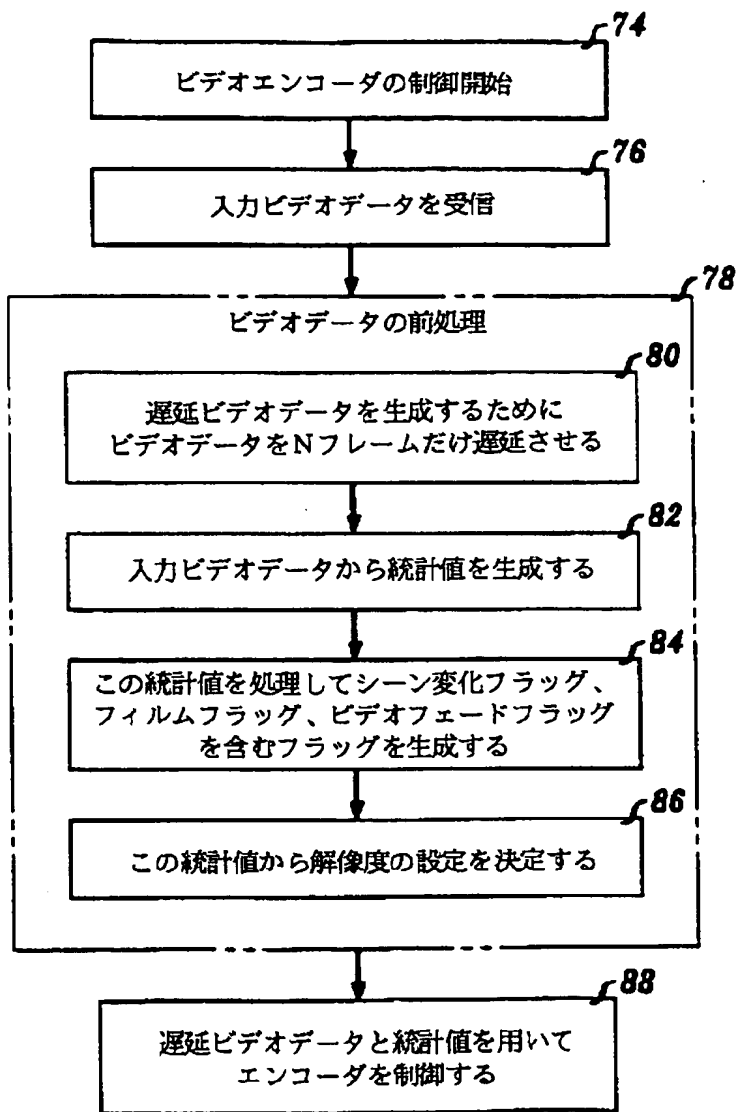
[Drawing 3]



[Drawing 4]



[Drawing 6]



[Translation done.]